Review of

Alaska Crab Overfishing Definitions

24-28 April 2006, Seattle, Washington

Report to

University of Miami Independent System for Peer Reviews

Michael C. Bell
Goose Cottage
4, Mobbs Cottages
Hall Lane, Oulton
Lowestoft
Suffolk, NR32 5DH
UK
bandm.bell@virgin.net

Contents

Executive summary

- This report is a review of proposed overfishing definitions (OFD) for Bering Sea and Aleutian Islands (BSAI) king and Tanner crab stocks. An OFD is required to meet National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act. The proposal is for a five tier system, specifying an MSY control rule within each tier, and is intended to replace the existing three tier system.
- The existing OFD provides no effective cap on exploitation rates. As a framework, the proposed OFD represents a major improvement. If successfully implemented it will meet the National Standard 1 requirement for MSY control rules which, if implemented as a harvest strategy, would be expected to result in a long-term average yield approximating MSY. There are, however, a number of issues that need to be addressed before the proposed OFD could be implemented.
- The proposed framework is comprehensive and adaptable, allowing the definition of MSY control rules in a very flexible way. The disadvantage of this flexibility is that it also implies complexity – there are a number of parameters for which default values will need to be determined before implementation.
- The main difficulty in establishing default parameter values, and in finding proxy values for reference points in Tiers 3 and 4 of the proposed OFD, is in the definition of effective spawning biomass (ESB). ESB is used in the MSY control rules and in the stock-recruitment relationships that are used to find proxy values for F_{MSY} and to test the performance of the proposed OFD under various parameterisations. Any satisfactory definition of ESB must (a) be demonstrably proportional to total fertilised egg production (TFEP), and (b) be responsive to fishing mortality. The first criterion is met by none of the definitions of ESB considered thus far. The problem arises out of the complex mating systems of king and Tanner crabs coupled with fisheries directed only at males. Simple mating ratios appear inadequate to capture this complexity.
- Suggestions for simple interim definitions of ESB are made in this report, together with recommendations for further research to identify more satisfactory alternatives to be used in the future.
- Simulation modelling was used to compare the performance of OFDs and to provide insight into likely default values for parameters in the proposed OFD. This approach is sound in principle and correctly specified in practice (in terms of model structures), but given that the simulation outcomes depend on a correctly specified measure for ESB no default parameter values can yet be recommended.
- Simulations were undertaken by two modelling teams. This is a strength in terms of allowing critical analysis of assumptions and robust conclusions. However, there are differences between the teams in their interpretation of the available scientific evidence on some fundamental issues of crab life history. These differences will need to be resolved in order to progress the simulation modelling to a final outcome.
- Simulation outcomes were largely judged in terms of rebuilding times for depleted stocks. Other aspects of OFD performance will need to be tested, such as the trade-off

between rebuilding times and level and constancy of yield. It should also be recognised that maintaining sustainable exploitation of healthy stocks is as important a function of an OFD as allowing recovery of depleted stocks.

- An important problem with the simulations was that the MSY control rule was treated as if it was a harvest control rule. This fails to recognise the role of the State in defining a precautionary buffer between target and limit fishing mortality rates. It is also likely to lead to selection of default parameter values for the proposed OFD that will place undue constraints on the capacity of the State to manage the fisheries according to precautionary and other objectives. It is recommended that MSY control rules are always tested in conjunction with realistic State harvest strategies in the simulations.
- It is concluded that, although work remains to be done before the proposed OFD can be implemented, the obstacles to successful implementation are not insurmountable. Given the urgent need for the existing OFD to be replaced by a more satisfactory alternative, it is recommended that a simple interim definition of ESB be adopted in the immediate term and that new simulations aimed at identifying default parameter values are undertaken at the earliest opportunity. These simulations will involve harvest strategies as well as MSY control rules.

Recommendations

Recommendations arising from this review are listed under *ToR (b): Recommendations of improvements to proposed overfishing definitions* (p.21) and *ToR (e): Suggested research priorities* (p.24).

Background

The North Pacific Fishery Management Council (NPFMC) has determined that the current overfishing definitions (OFD) for Bering Sea and Aleutian Islands (BSAI) crab stocks are in need of revision. Proposals for revised OFDs have been developed by a four member Work Group reporting to the Crab Plan Team (CPT). A panel of three independent reviewers was invited by the Alaska Fisheries Science Center (AFSC) to review these proposals, along with simulation models used to test their performance and to determine default parameter values and proxies.

The review panel members were Patrick Cordue (independent consultant, New Zealand), Nick Caputi (Department of Fisheries, Western Australia) and the present author (Michael Bell, independent consultant, UK). The Terms of Reference for the review were:

(a) A statement of the strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.

(b) Recommendations for improvements to proposed overfishing definitions or alternative definitions,

(c) A review of the model configurations, formulations and methods used to account for uncertainty.

(d) A review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in simulation models.

(e) Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

The summary of review findings given below is structured according to these Terms of Reference, although overlaps in the relevance of items means that most of the issues are covered under ToR (a). This report represents the individual opinion of the present author. No attempt was made to reach a consensus among the three reviewers, but it was apparent during the review meeting that differences among the reviewers are likely to be in emphasis rather than substance.

Description of review activities

Documents relating to overfishing definitions and management of Bering Sea and Aleutian Islands (BSAI) crab stocks were provided to reviewers on the web site www.afsc.noaa.gov/refm/stocks/CrabWs.htm. This web site was initially developed as part of an inter-agency workshop on crab overfishing definitions held in February 2006 in preparation for the CIE review and NPFMC action. Appendix 1 lists the key documents on this web site and other documents provided during and after the meeting. Prior to the meeting attention was drawn to a number of key documents which provide the necessary background for the review meeting:

(1) the Statement of Work for the Work Group responsible for developing proposals for the overfishing definition (Rugolo, 2004);

(2) a description of the proposed tier system for the overfishing definition (NPFMC, 2006; Siddeek & Zheng, 2006; Turnock & Rugolo, 2006a);

(3) stock assessments for Bristol Bay red king crab (Zheng, 2004) and eastern Bering Sea snow crab (Turnock & Rugolo, 2005);

(4) position papers discussing unresolved issues for the Work Group (Turnock & Rugolo, 2006b; Zheng, 2006)

(5) report and recommendations from the February workshop (NPFMC, 2006); and

(6) results of projections examining the performance of the proposed overfishing definition (Siddeek & Zheng, 2006; Turnock & Rugolo, 2006a).

A review meeting took place at the Alaska Fisheries Science Center, Seattle, 24-28 April 2006 (see Agenda at Appendix 2). The meeting was chaired by Anne Hollowed and Jim Ianelli of the NMFS. The meeting was introduced by Anne Hollowed, followed by a description of crab management and the need for a revised overfishing definition by Diana Stram, NPFMC. Over the course of three days, members of the interagency Work Group charged with developing proposals for a revised OFD (Shareef Siddeek and Jie Zheng of ADF&G, Jack Turnock and Lou Rugolo of NMFS) presented overviews of the proposed OFD and tier system, assessments for snow crab in the eastern Bering Sea and red king crab in Bristol Bay, approaches to estimating proxy values for biological reference points and simulations testing the performance of OFDs. Extensive discussions with CIE panel members took place alongside the presentations, so that this part of the programme extended to the end of the third day of the meeting. CIE panel members met on day 4 to discuss the main issues raised during the presentations, and sought some clarifications from NMFS staff involved in the review meeting. The remainder of the review meeting time was spent in preparing to write individual review reports.

Summary of findings

General

The existing OFD for BSAI crab stocks consists of three tiers, from Tier 1 for stocks with the least amount of information on stock status and exploitation to Tier 3 for stocks with the most amount of information (Turnock & Rugolo, 2006a). The maximum fishing mortality threshold (MFMT) is set to F_{MSY}, assumed to be equal to M (set to 0.2 for king crabs and 0.3 for snow and Tanner crabs). The minimum stock size threshold (MSST) is set to $½B_{MSY}$ for stocks in Tier 3, where B_{MSY} is assumed to take the value of the average of survey estimates of mature male and female biomass during 1983-97. MSST is undefined for stocks in Tiers 1 and 2. MSY is determined either as the product of F_{MSY} and B_{MSY} (Tier 3) or from a proxy of mature biomass and stock utilisation rate (Tiers 1 and 2).

As pointed out in a presentation on the Statement of Work for the CPT Work Group (Rugolo, 2006), the existing OFD is unsatisfactory in a number of respects. Most importantly, the definition of sustainable yield involves all mature crabs, both sexes and all shell classes, irrespective of vulnerability to the directed fishery, and provides no effective cap on exploitation rates. Rugolo (2006) provides an example where catch levels for snow crabs could be set higher than the total exploitable biomass (legal males), without overfishing being declared. Projection models used by Turnock & Rugolo (2006a) demonstrate that fishery management under the current OFD cannot provide effective rebuilding to B_{MSY} from overfished stock levels for both red king and snow crab stocks (notwithstanding concerns about definitions of effective spawning biomass in these simulations – see below, p.15). The need for a revised OFD is very clearly established. The review findings presented below indicate that much work remains to be done before a revised OFD could be accepted, but retaining the *status quo* OFD is not a tenable option for the immediate future.

ToR (a): Strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches

Framework for overfishing definitions

Proposals for a revised OFD for BSAI crab stocks involve a system of five tiers, from Tier 1 for stocks with the most complete and reliable assessments to Tier 5 for stocks with data only on the catch history (Siddeek & Zheng, 2006; Turnock & Rugolo, 2006a). An MSY control rule for Tiers 1 to 4 involves defining F_{MSY} and B_{MSY} or proxies, and calculating the overfishing limit for fishing mortality F_{OFL} in terms of these values and parameters which define the slope of F_{OFL} in relation to stock biomass and the threshold stock level below which the fishery is closed. Figure 1 shows this proposed MSY control rule compared with the default MSY control rule advocated by Restrepo *et al.* (1998) in technical guidance on implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act (MSFCMA). The proposed OFD is much more conservative than the default OFD in that for any positive value of the parameter α the value of F_{OFL} will always be lower for any given stock biomass level. The suggested default for MSST is $(1–M)*B_{MSY}$, which accords with the notion that under MSY harvest levels the scale of fluctuations in biomass around B_{MSY} is likely to be in the order of $M*B_{MSY}$. Although the proposed setting of MSST at $½B_{MSY}$ is lower than this default, this is scarcely relevant since reductions in fishing mortality at higher biomass levels provide for returning stock trajectories towards B_{MSY} even in the absence of special stock rebuilding plans. Indeed, it could be argued that MSST could be dispensed with altogether under the new proposals, although there is certainly merit in having a trigger point at which the effectiveness of F_{OFL} levels under the control rule are re-examined.

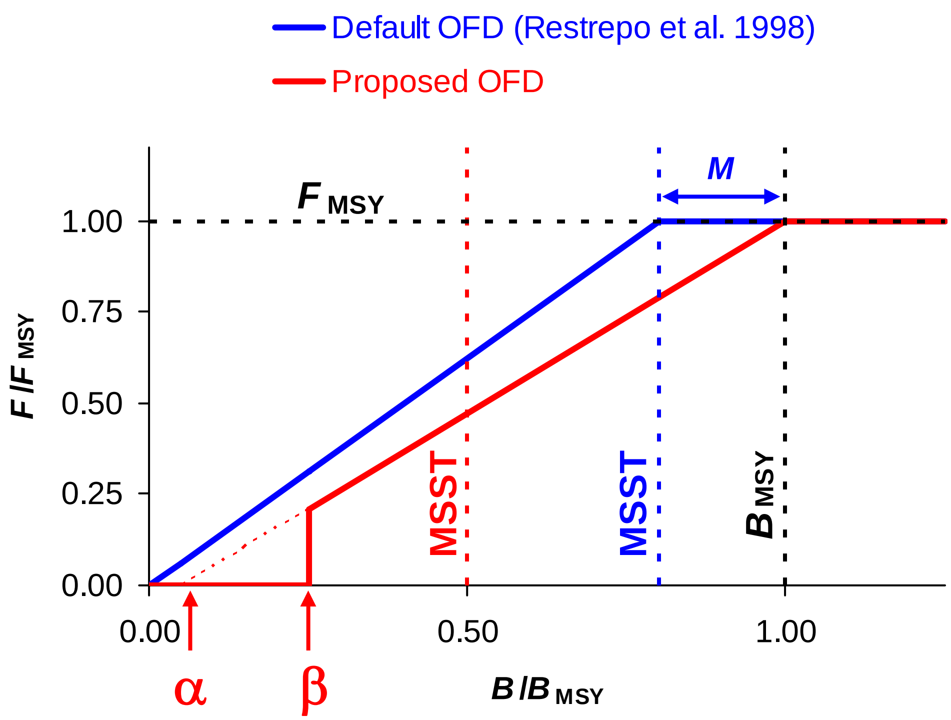

FIGURE 1. Proposed MSY control rule (red lines and captions) shown in relation to the default MSY control rule (blue lines and captions) put forward in technical guidance by Restrepo *et al.* (1998).

As a framework, the proposed OFD represents a major improvement over the existing OFD. National Standard 1 of the MSFCMA requires an MSY control rule which, if implemented as a harvest strategy, would be expected to result in a long-term average catch approximating MSY. If (and only if) successfully implemented, the proposed OFD would be expected to fulfil this requirement in the sense that the capacity of stocks to support harvests up to MSY should not be compromised by excessive fishing mortality. The same certainly could not be said of the existing OFD. As described below, there are a number of issues which need to be addressed before the proposed OFD could successfully be implemented. Given the inadequacy of the current OFD, it is vital that issues related to the implementation of the proposed OFD be resolved. Emphasis is given below to short-term actions that could be progressed on a time-scale to allow early implementation of the proposed OFD.

A second major strength of the proposed framework is that it is comprehensive. This is true in two senses. First, the use of a five tier system allows account to be taken of the state of knowledge of the stock and the reliability of assessments and monitoring data. Siddeek & Zheng (2006) and Turnock & Rugolo (2006a) allocate most of the 22 BSAI crab stocks to Tier 5, for which the overfishing limit depends only on an average of historical catches. No stock is expected to be allocated to Tiers 1 and 2, which require at least point estimates for F_{MSY} and $B_{MSY,}$ and only three stocks are allocated to Tier 3, requiring a reliable proxy for F_{MSY}. As surveys, assessments and monitoring systems improve, and as improved estimates of biological parameters become available, it would be expected that stocks could be promoted within the tier system. This perhaps applies mostly to Tier 3 and 4 stocks; information is perhaps likely to remain scanty for Tier 5 stocks taken mainly as a by-catch in groundfish-directed fisheries.

Second, the framework is comprehensive in the sense that it has several parameters which allow the dependence of F_{OFL} on stock biomass to be defined in a very flexible way. The α parameter acts as an *x*-intercept on the MSY control rule graph, determining how quickly F_{OFL} is reduced as biomass decreases, while the β parameter sets a biomass threshold for closure of the fishery (Figure 1). This allows great flexibility in defining the MSY control rule: *e.g.* $\alpha=-\infty$, $\beta=0$ defines a flat F_{MSY} control rule; $\alpha=\beta$ allows F_{OFL} to take any value between 0 and F_{MSY}, depending on biomass; *etc*. This flexibility allows the capacity for evolution, adaptation and refinement in the implementation of the OFD for individual stocks as more information becomes available.

There are also disadvantages to this flexibility. In the first place, freedom in defining the shape of the MSY control rule presents challenges for setting up starting defaults. Parameters α and β are arbitrary, in the sense that they have no objective definition – *e.g.* β is not defined as the threshold biomass at which there is an *x*% probability of event *y* occurring. This in itself is not necessarily a drawback, since it is the operational properties of the parameters that we are interested in, but it does mean that we can only determine the best combinations of parameters by examining the emergent properties of the systems in which they are defined. This requires extensive simulations (as by Siddeek & Zheng, 2006), in which we need to determine the criteria by which we judge one outcome better than another (see below).

A second disadvantage of the flexible formulation for the MSY control rule is that capacity for evolution also implies being subject to change. Given revision of assessments on an annual basis, there is the capacity to revise the parameters of the OFD on each occasion that it is applied. Revision of biomass and fishing mortality estimates as new data are added inevitably will change the perception of past stock status: biomass and fishing mortality levels previously considered as being within precautionary limits might, in the light of new data, be considered as representing overfished or overfishing states, and *vice versa*. This in itself is not a problem, as we can only act in the present, based on the best available information on current conditions. What is a problem, however, is that it is not just the estimates themselves that change, but also the reference points against which they are compared. This applies to F_{MSY}, B_{MSY} and their proxies (*i.e.* Tiers 1 to 3), which depend on the nature of the stock-recruitment relationship (SRR). Given poorly defined SRRs (*e.g.* Zheng, 2004), there is the capacity for new data points to be highly influential. Values α and β are less subject to annual change, given that they are selected rather than estimated, but it should be remembered that their influence on the performance of the OFD is inextricably linked with the particular values taken by F_{MSY} and B_{MSY}. For stocks close to overfished thresholds, even minor changes in the OFD could lead to instability in the management regime – *e.g.* opening and closing of the fishery. Radical changes in perception, if accepted as plausible, of course require radical management responses. Otherwise protocols are required for stabilising the OFD in the face of changing assessments. One approach would be to use moving averages (*e.g.* of F_{MSY} and B_{MSY} proxies in Tier 3) to reduce annual biases (as suggested by Nick Caputi during the review meeting). Another approach would be to set up a cycle of regular update assessments, where many assessment parameters would remain unchanged, and occasional full assessments with revision of F_{MSY} and B_{MSY} or proxies and testing whether current values of α and β remain appropriate. ‘Update’ assessments would be the sole responsibility of assessment authors, whereas ‘full’ assessments would have consequences for rigorous documentation and Council review. The latter approach is similar to that adopted by ICES in recent years.

The two most important components of the OFD are stock biomass and fishing mortality, *i.e.* the *x*- and *y*-axes for the MSY control rule (Figure 1). The way these two parameters are defined is critical to the successful operation of the OFD. In the existing OFD there are logical inconsistencies between the way these are defined in the threshold values and the way they are applied in determining harvest levels. The proposed OFD potentially resolves these inconsistencies, but successful implementation of the proposed framework depends critically on the correct definitions of biomass and fishing mortality. The definition of fishing mortality appears not to be a concern. As defined in the framework, F applies to all vulnerable portions of the population, including discards of females and undersized/unmarketable males. Mortality of trawl by-catch is a separate issue (see below), but provided that this is adequately accounted for in the assessment and simulation models, this is not a problem in terms of the framework. Within the framework, ‘F’ refers to δF for fully selected crabs, where δ is the time interval over which the fishery occurs. Provided that F_{MSY} and F_{OFL} are expressed in this same currency, then the way that fishing mortality is defined in the OFD framework is satisfactory. The same cannot be said of stock biomass. ‘Biomass’ here refers to spawning potential, and in fact need not be expressed in biomass units at all. As emphasised by Patrick Cordue on numerous occasions during the review meeting, the biomass measure would be expected to be proportional to total fertilized egg production (TFEP). Discussion of this critical issue in relation to MSY control rules and SRRs is deferred to a later section (see p.15). It is enough to note here that the various options for defining effective spawning biomass (ESB) considered in the simulation studies (Siddeek & Zheng, 2006; Turnock & Rugolo, 2006a) appear unlikely to meet this criterion of proportionality with TFEP. Successful implementation of the proposed OFD will not be possible until a satisfactory definition of ESB is determined. This issue affects the definition of SRRs on which estimation of F_{MSY} or proxies and testing of the OFD depends.

Estimating proxy values for biological reference points

In 'data-rich' situations (stocks in Tiers 1 and 2), F_{MSY}, B_{MSY} and associated quantities can be estimated directly, but for all other cases we need proxies for these values. Restrepo *et al.* (1998) offer extensive guidance on selecting among the various candidates for proxies in 'data-moderate' and 'data-poor' situations. Siddeek & Zheng (2006) and Turnock & Rugolo (2006a) opted to consider $F_{x\%}$ values as proxies for F_{MSY} for Tier 3 stocks, where x is the percentage of virgin spawning potential per recruit (SPR) at equilibrium. Given that the OFD requires a working definition of spawning potential (*i.e.* ESB) irrespective of the approach to deriving proxies, selection of proxy values on the basis of SPR is a sensible approach despite the difficulty in defining ESB. Restrepo *et al.* (1998) advocate the use of $F_{x\%}$ in preference to yield per recruit reference points ($F_{0.1}$ and F_{max}) and SRR-based reference points (F_{med}). NPFMC (2006) recommended the range $F_{50\%}$ to $F_{60\%}$ based on previous work by the CPT Work Group. According to the method of Clark (1991), whereby the most likely value of F_{MSY} is selected at the intersection of yield curves from the most and least productive of a plausible range of SRRs, this range of $F_{x\%}$ appears reasonable. At present, the use of $F_{x\%}$ and the approach to selecting the appropriate x% can both be endorsed as satisfactory approaches to deriving F_{MSY} proxies for Tier 3 stocks, but the issue cannot reasonably be progressed further towards selection of actual values without first resolving issues relating to the definition of ESB and hence SRRs (see p.15). Two further points can be made in this context. First, SPR-based reference points are likely to be highly sensitive to assumptions about growth. Estimation of growth patterns was not discussed in detail during the review meeting, but it seems safe to suppose that moult frequencies and increments are fairly poorly resolved for BSAI crab stocks. New information on growth potentially could be very influential in identifying reference points. It is recommended that the sensitivity of reference points and the performance of the OFD be explored in relation to uncertainty about growth. Second, the question of the likely form of the SRR should be addressed. For example, are there *a priori* reasons for selecting a Ricker rather than a Beverton-Holt SRR, *e.g.* likely cannibalism of pre-recruits by the adult stock? It is recommended that the plausible range of SRR types be examined carefully in the light of what is known about recruitment biology, with a view to constraining the range of SRRs that are considered in selecting biological reference points.

For Tier 4 stocks the proxy for F_{MSY} is defined as γM. The use of M as the basis for a proxy is consistent with the guidelines given by Restrepo *et al.* (1998) for 'data-poor' situations, and appears sensible given the available options. However, determining default values for γ presents difficulties. Presumably, values would be selected at the group- rather than stock-level, where one group was king crabs (*Lithodes* and *Paralithodes* spp.) and the other was snow and Tanner crabs (*Chionoecetes* spp.). Additional information on individual stocks seems likely to result in promotion to Tier 3 rather than modification of γ. The precise role of γ in the proxy needs to be clarified. It appears that it is not involved in a 'currency' change in the OFD, given that fishing mortality is already expressed in terms that account for the timing and selectivity of the fishery (see above). Instead, γ appears genuinely to scale M towards the appropriate F_{MSY}. In the absence of further information, it seems reasonable to suppose that γ=1 would be an appropriate default, but analyses and simulations of the type applied to Tier 3 stocks may be informative about the most likely values.

Simulation modelling

Simulation models were separately developed by ADF&G staff (Siddeek & Zheng, 2006) and NMFS staff (Turnock & Rugolo, 2006). Differences of approach potentially shed light on the sensitivity of simulation outcomes to particular modelling choices, and final conclusions about the performance and parameterization of the OFD could be more robust as a result. There is a need for critical analyses comparing the results of the two simulation approaches, but this will first need agreements on some critical biological issues and some common grounds for comparison (see below).

Simulation models for BSAI crabs are likely to have much in common with the length-based assessment (LBA) models: the first are used for forward projections, the second to estimate parameters. If assessment parameters (both estimated parameters and those fixed *a priori*) are to be used in simulation models, it is vital that the model structures be identical, since they are likely to be valid only in the context in which they are estimated or applied. Modelling of uncertainty in the population dynamic and survey processes should preserve the covariance structure of parameters estimated in the assessments.

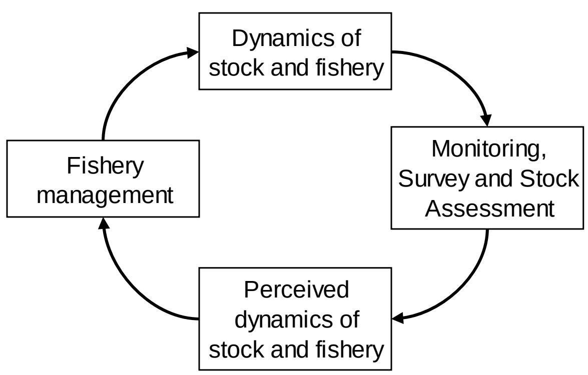

FIGURE 2. Processes involved in simulating fishery management.

As noted by Restrepo *et al.* (1998), various sources of uncertainty are involved in modelling fishery management. These include uncertainty owing to the accuracy and precision of estimates, choice of model structure, natural variability of stock dynamics and errors in the implementation of management measures. Figure 2 shows schematically the major processes involved in simulating a fishery management system – each box and each arrow comes with its own component of uncertainty which should be incorporated within the simulation. The distinction between the 'perceived' and 'true' dynamics is particularly important. Turnock & Rugolo (2006a) incorporate observation errors by applying autocorrelated lognormal errors to abundance within the simulations. This approach is probably adequate for the purposes of comparing the performance of alternative OFDs, but it is recommended that in the long-term the model be extended to include the full assessment-management processes.

An important issue for the simulations undertaken by both Siddeek & Zheng (2006) and Turnock & Rugolo (2006a) is that they treat the MSY control rule as if it were both a harvest strategy and a rebuilding plan. Fishing mortality in the simulations always takes the value of F_{OFL}. In one sense this could be seen as fair enough, since this is a worst case scenario and the OFD is already very precautionary as a framework and could conceivably be viewed as constituting a default rebuilding plan for depleted stocks. Restrepo *et al.* (1998) note that an MSY control rule that incorporated 'built-in' rebuilding might be used "if a Council wished to minimize the range of stock sizes within which special rebuilding plans would be required". However, if it really is the case that the MSY control rule is seen as sufficiently precautionary to be used as a harvest strategy, then either it implies that the OFD is likely to be unduly restrictive in that it allows very little room for manoeuvre by the State fishery managers, or else it appears to by-pass the requirement for the State to act in a precautionary manner in maintaining a buffer between F_{OFL} and the target F used in setting TACs. At the State level, harvest strategies may well incorporate multiple management objectives, going beyond mere stock conservation. Indeed, for fisheries that are economically as well as biologically healthy, it is desirable that management objectives incorporating socio-economic aims should explicitly be stated. An MSY control rule that seeks to take the role of a default harvest strategy is unlikely to be conducive to such enlightened management. This is relevant to the current simulations, because it implies that they are modelling scenarios that will never (or at least should never) happen in practice. The role of the OFD is not to replace the requirement for precautionary management by the State, but to provide the context in which this can occur. State management could either use the OFD as a Federal check on the admissibility of their preferred harvest strategy, or as a fixed point of reference to determine their harvest strategy (*e.g.* setting TACs consistent with 0.75 F_{OFL}). In either case, it is the harvest strategy, not the MSY control rule, that is applied to the stock. It is therefore recommended that a credible State harvest strategy is *always* included in simulations of the performance of an OFD for BSAI crab stocks. If, on the other hand, it *is* intended that the OFD should take on the role of a precautionary harvest strategy, this should be explicitly stated and the performance of the OFD should be considered in terms beyond average stock size and rebuilding times.

It is in any case desirable to examine multiple aspects of the performance of a proposed OFD. It is relatively easy to define management measures that are just precautionary, less easy to define ones that balance precaution against other objectives for a fishery. For depleted stocks, there is an obvious trade-off between short-term pain and long-term gain. In other words, shorter rebuild times to higher stock levels come at the expense of immediate losses of yield. More generally, management responses to changing stock sizes have consequences for the level and, particularly, the variability of yield. Different

OFDs with similar properties in terms of rebuild time and probability, may differ strongly in their properties with respect to short-term losses and long-term constancy of yield. It is recommended that trade-offs involving yield (or any other fishery management objectives) be included in simulation studies of the performance of OFDs for BSAI crabs. Again, it should be emphasised that simulations should not consider OFDs in isolation from harvest strategies.

Even if we set aside management objectives other than stock conservation, an OFD can be viewed as serving two roles: (i) it is intended to allow depleted stocks to grow towards biomass levels capable, on average, of supporting MSY; and (ii) it is intended to prevent the fishery from causing stock biomass to decline below these levels. The simulations for BSAI crabs have focussed on the first role to the exclusion of the second. In part this is natural, since the α and particularly β parameters are most relevant to recovery from low stock sizes. However, recognising the role of the OFD in allowing fisheries to operate at sustainable levels over the long-term, it is recommended that the simulations use a variety of different starting biomass levels up to B_{MSY}, rather than just considering depleted stocks. Note that, for the purposes of comparing between the Siddeek & Zheng (2006) and the Turnock & Rugolo (2006) modelling approaches, the same selection of starting biomass levels (in terms of fractions of B_{MSY}) should be used by both modelling teams.

The biggest differences between the two modelling teams were in their interpretation of certain biological issues critical to the definition of the spawning stocks. Many issues that divide the two groups are highlighted in the ADF&G and NMFS position documents (Turnock & Rugolo, 2006b; Zheng, 2006). The February Workshop Report (NPFMC, 2006) makes what are intended to be definitive statements on some of these issues, but it is apparent that it is possible to interpret these statements in more than one way. For example, at the review meeting it was particularly apparent that different views had been taken about the minimum interval between moulting and mating for new shell male snow crabs, with strong implications for their participation in primiparous and multiparous matings and hence for the definition of the male spawning stock. It is beyond the remit of the present review to arbitrate on such issues, but there is a strong need either for the two modelling teams to agree on the interpretation of the best available scientific information, or for this interpretation to be determined by a third party. In cases of genuine uncertainty about the biological processes, this should be included in the simulations as sensitivity analyses.

Effective spawning biomass and the stock-recruitment relationship

If there is one crucial issue on which the proposed OFD succeeds or fails, it is in the definition of effective spawning biomass (ESB). ESB plays two roles in the OFD: first, it is the *x*-axis of the MSY control rule, determining the value of F_{OFL} and being the scale on which MSST and B_{MSY} are measured; second, it is the controlling variable for the stock-recruitment relationship (SRR), used in determining F_{MSY} or its proxy and in testing the performance of the OFD. The two roles are linked, since the outcome of applying the MSY control rule is intended to be a long-term average catch approximating

MSY (Restrepo *et al.*, 1998), and this outcome is achieved through translation of ESB into future recruits. To clarify: the MSY control rule does not simply avoid the recruitment failure that ultimately leads to stock extinction, although successful operation of the rule should in fact achieve this aim; rather, it has the more positive aim of encouraging stock size to reach a level that maximizes the delivery of biomass to the directed fishery. The first is more characteristic of the ICES paradigm of precautionary fishery management, and requires knowledge or assumptions about the left-hand portion of the SRR, *i.e.* what happens to recruitment at low stock sizes. The second, which applies under the National Standard Guidelines, requires knowledge or assumptions about the full form of the SRR, *i.e.* what happens at all stock sizes.

As noted above, it was agreed at the review meeting that the first essential property of any measure of ESB is that it should be proportional to total fertilised egg production (TFEP). A secondary property, needed for successful definition of an OFD, is that ESB should be sensitive to fishing mortality. Indeed, if ESB was not responsive to changes in fishing mortality, there would be no point in having an OFD! These two requirements have led to some widely varying definitions of ESB by the two Work Group teams (Siddeek & Zheng, 2006; Turnock & Rugolo, 2006a). The most obvious definition for ESB is the total biomass of mature females. This falls at both hurdles – the first because TFEP depends also on the availability of males as mating partners, the second because fishing mortality is directed at males, females being sensitive only to by-catch and discard mortality. The next step is to suppose that each male present in the population can mate with a certain number of females (the ‘mating ratio’) and that ESB is best defined by the minimum of total biomass of mature females and the biomass of mature females that is capable of being mated by the mature males present in the population. Setting aside the issue of determining which males participate in mating (see above), this is an improvement in that it recognises that TFEP is limited by the availability of males. However, the use of a mating ratio is a gross simplification of the complex mating system of snow, Tanner and king crabs, and this definition of ESB is still only weakly sensitive to fishing mortality. Turnock & Rugolo (2006a) attempted to deal with the latter issue by adding in the total biomass of mature males. This, however, is logical only as an *ad hoc* measure, and takes a step further away from a definition of ESB that is acceptable on biological grounds. Siddeek & Zheng (2006) used the mating ratio to calculate a male component to ESB in complement to the female component. Although an improvement, this still falls a long way short of a biologically realistic definition that could be used with confidence to test, parameterise and apply a working OFD. Estimation of F_{MSY} and proxies and other outcomes of simulations were found to be very sensitive to mating ratio and other facets of the definition of ESB.

The central problem is that a simple, robust and biologically meaningful index of TFEP is required, whereas the complexities of BSAI crab mating systems allow no simple answers. Sperm storage by females, sperm rationing by males, size-assortative mating, non-participation in mating by new shell males, lower clutch fullness in primiparous females, the existence of ‘graveyard’ females – these and, no doubt, many other factors make it extremely difficult *a priori* to write down a suitable expression for ESB based on simple, easily measurable quantities. Given size-fecundity relationships that appear to be linear rather than cubic functions of carapace width in snow crabs (within clutch fullness categories), it is even questionable whether spawning potential is best measured as a biomass. Patrick Cordue suggested during the review meeting that it would be enlightening to construct an individual-based model (IBM), incorporating and simulating the various features of crab mating and egg production systems. Hypotheses about the causes and consequences of size-assortative mating may profitably be explored in this context – changes in the rate of successful mating as population density changes are likely to differ strongly according to whether male-male competition (ousting of small males by large males) or loading constraints (limits to the size of female that can be handled by a male of a given size) are the primary cause of size-assortation. Spatial factors, in relation to migration patterns and the location of primiparous and multiparous matings, could also be explored using an IBM. Imperfect knowledge of some or even many aspects of these systems may make it difficult to parameterise an IBM with any certainty. However, given plausible assumptions it may be possible to draw deductions about at least the functional form of any satisfactory definition of ESB. Furthermore, an IBM may provide insights into the conditions under which the effective sex ratio may be genuinely limiting for egg production. Although the primary aim of the OFD is to allow optimum recruitment rather than to prevent recruitment failure, it is nevertheless useful to know the circumstances under which this might be expected to occur.

A second approach to finding an appropriate functional form for ESB would be to examine data that are already available from the annual surveys. In some years at least, when environmental temperatures favour the presence of ovigerous females in the population at the time of the survey, it may be possible to measure egg production directly. Data on female size and clutch fullness are routinely collected, which would allow calculation of total egg production. Calculation of TFEP would depend on whether or not it is possible to draw deductions about fertilisation rates from egg colour (blue coloration means a developing and hence definitely fertilised egg, orange coloration could mean either unfertilised or simply early stage). Is it perhaps possible to make deductions about relative fertilisation rates, even if there is no confidence that these can absolutely be estimated? If so, then both sides of the equation relating TFEP to crab population structure are known in at least relative terms – it merely (!) remains to estimate the functional form.

The IBM and the survey data approaches to estimating ESB, or at least a functional form for ESB, are both medium- to long-term projects, and are thus unlikely to yield results that are useful for the timely implementation of the proposed OFD. A short-term, probably interim, solution is required. Replacement of the existing OFD is an urgent priority, even if this requires a less than perfect solution to the definition of ESB. Given fisheries that preferentially remove males from their target populations, we know that male availability is the most likely limiting factor for successful reproduction in BSAI crabs. Accordingly, any meaningful definition of ESB must include mature males. Perhaps the most likely candidate for an interim definition of ESB is the total mature male biomass. This definition was proposed at the review meeting by Nick Caputi. It has the virtues of being simple (given agreement as to what constitutes the mature male population) and being responsive to fishing mortality. This definition will certainly be *wrong*, particularly at higher stock sizes where it is more likely that recruitment will be egg-limited than sperm-limited[1], but it has the highly desirable property that it may be a very good measure of the degree to which spawning potential is impaired at low stock sizes, at which the MSY control rule will cause fishing mortality to be reduced or the fishery to be closed. As noted above, this does not accord entirely with the philosophy of the OFD, but it is nevertheless an extremely important function for it to fulfil. Moreover, if at higher biomass levels stock size assumes a relatively minor role compared with environmental factors (see below), *i.e.* 'noise' makes a greater contribution to recruitment variation than the underlying SRR, then an incorrectly specified ESB is much less of a problem.

The following course of action is recommended for incorporating a stock biomass measure into the OFD and in modelling SRRs for BSAI crab stocks:

(1) Agree a simple short-term interim measure for ESB and prescribe that this is the definition that will be used in implementing the proposed OFD. Total mature male biomass is a strong candidate for this measure.

(2) Again in the short-term, use estimates from the stock assessments to explore the relationship between recruitment and the two axes of mature male biomass and mature female biomass. It is unlikely that data will be available over much of this surface, but it is worth establishing the extent to which variability in recruitment can be accounted for by the joint effects of these two variables.

(3) In the medium-term, examine the suitability of survey data for estimating TFEP, and explore the possibility of using these data to determine the appropriate functional form for ESB. If necessary, this should be supplemented with direct field measurements of clutch fertilisation rates.

(4) In the medium- to long-term, construct an IBM of BSAI crab mating and egg production systems, aiming to determine the appropriate functional form for ESB. It is important that the IBM be spatially structured, to examine the spatial co-incidence of different population components at mating time and to consider the delivery of larvae to suitable settlement areas.

Before leaving the topic of SRRs it is worth considering sources of variation in recruitment other than ESB. This is relevant in two respects. First, if ESB plays a relatively minor role in determining recruitment, this is important both for simulating the performance of the OFD and for approaches to management. For recruitment-driven fisheries, conservation of spawning stock biomass becomes less important compared with managing the mortality of recruits that are delivered to the fishery (subject, of course, to precautionary minimum stock biomass levels). This is particularly so if natural mortality is high and recruitment is very variable between years. This is perhaps more typical of

[1] In fact the application of a mating ratio to mature female biomass could be seen as dealing with the transition from sperm-limitation to egg-limitation as stock size increases, but the sensitivity of F_{MSY} to choice of mating ratio together with other uncertainties make it preferable to adopt a more simple approach. The use of mature male biomass will introduce its own problems for selection of an F_{MSY} proxy, but choice of a value between $F_{50\%}$ and $F_{60\%}$ could be made based on operational properties even if it could not be defended as an unbiased proxy.

shorter-lived species than snow, Tanner or king crabs, such as estuarine bivalves or blue crabs, but it might be worth considering the effects of such recruitment patterns in the simulation models (some simulations of random recruitment were shown by Shareef Siddeek during the review meeting). This also highlights the importance of ensuring that the error components of the SRRs in the simulations are adequately characterised.

The second point of relevance is the issue of regime shift and stock productivity. Changes in abundance and distribution of some BSAI crab stocks have been attributed to changes in the climatic and oceanographic regime in the Bering Sea, although a lack of coherence of change between different stocks may cast some doubt on this hypothesis (Zheng & Kruse, 2006). Dew & McConnaughey (2005) emphasise the role of fishing mortality rather than regime shift in causing the decline of Bristol Bay red king crab after 1980, suggesting *inter alia* that intensive trawling in an area important to spawning females may have affected delivery of larvae to the most suitable grounds for settlement. Deciding whether or not regime shift is responsible for recent recruitment levels is extremely important, given that management priorities for a depleted stock will be very different to those for a stock which has simply become less productive. Restrepo *et al.* (1998) note that "for a period of declining abundance, the 'burden of proof' should initially rest on demonstrating that the environment (as opposed to fishing) caused the decline, and that, therefore, the target control rule should be modified". It is beyond the scope of the current review to offer a view on which are the most likely causes of changes in abundance of BSAI crab stocks, but it is recommended that a consensus be sought among the relevant experts about whether regime shift or fishing mortality are the most likely causes of change, and that this consensus be used to inform development of the most appropriate OFD.

Stock assessments

Stock assessments for Bristol Bay red king crab (Zheng, 2004) and eastern Bering Sea snow crab (Turnock & Rugolo, 2005) were briefly presented at the meeting. In both cases a length-based assessment (LBA) model was constructed, accounting for the complex, discontinuous growth patterns of the two species. Given the emphasis on other issues, it was not possible to examine the assessments in detail during the meeting. However, the assessments are essential to the OFDs because: (i) they yield values which will eventually be compared with the criteria of the MSY control rules; and (ii) they yield values which allow the OFD to be parameterised and tested. Brief comments on some of the main features of the assessments are given below, but this should not be treated as a full assessment of the assessment methodologies and outcomes. A review of the eastern Bering Sea snow crab assessment has recently been undertaken on behalf of CIE by Maunder (2003).

LBA models are used for both red king crab and snow crab assessments. The two approaches differ principally in the number of free parameters (more for snow crabs) and method of estimation (sum-of-squares based on lognormal errors for red king crabs, maximum likelihood for snow crabs). The method of dealing with size transitions in each model appears satisfactory, if based on rather slender data resources. Lack of opportunity to examine this in detail prevents further comment on this potentially important issue, but the dependence on and sensitivity to poorly known biological parameters in the assessments warrants further study in the future. Given the complexity of these models – in terms of both structure and number of parameters – it would be reassuring to see corroboration of the broad patterns of assessment results drawn from comparisons with the outcomes of simpler assessment approaches (*e.g.* Collie-Sissenwine Analyses) which make fewer demands on knowledge of biological parameters such as growth.

In general, biological realism is a strength in these LBA models (to the extent that this realism can be supported by available information), but within the biologically realistic model structures it is desirable to seek the most parsimonious descriptions (fewest parameters) of the data and processes. In the case of the snow crab model almost 300 parameters are estimated and it is appropriate to ask whether every one of these is necessary, or indeed supported by the data. In particular, selectivity parameters proliferate, with variation according to years, sexes and shell conditions; it seems likely that this complexity could be drastically reduced by applying a few well chosen *a priori* assumptions about selectivity patterns, with advantages for the precision of estimates. Some parameter sharing may also be possible within models. For example, in the red king crab model is it reasonable to suppose that patterns of annual recruitment are likely to be similar between males and females? If so, there is no need to estimate the two patterns separately; any departures from 1:1 sex ratio at recruitment can be dealt with by including an additive parameter rather than treating the patterns as completely independent. In general, it is recommended that maximum use be made of external information, whether as potential explanatory variables (*e.g.* environmental signals) for recruitment patterns or to determine the likely values of parameters to be estimated.

Both assessments are tied in to survey data, with the assumption that the survey represents the complete population. This assumption is precautionary, since it will tend to give an upwards bias to estimates of F and a downwards bias to estimates of population abundance. However, the assumption is unnecessary since the survey catchability could be estimated within the assessment: it is recommended that estimation of survey q be included within the assessments. It is further recommended that depletion experiments be undertaken using the survey gear and vessel to assess gear efficiency and selectivity. This approach is preferable to the ‘underbag’ experiments which have so far been used.

The red king crab assessment model includes estimation of M values for four different periods in females and for three different periods in males. This is done primarily to account for differences in population dynamics during the early 1980s. It is not clear that there is an objective basis for selection of the periods (which differ between males and females) beyond achieving a closer fit to the survey data. Without data on changes in specific mortality factors, it is not defensible to use *ad hoc* model adjustments to infer changes in M on the basis of model fit. Lack-of-fit could be due to factors other than natural mortality, such as catchability changes or changes in the relative spatial distributions of the stock and the fishery. Furthermore, ‘M’ as estimated in the red king crab assessments is actually a compound of natural mortality and indirect (by-catch) fishing mortality. It is highly desirable to separate these components, perhaps using effort data from the by-catch fleets.

More use of fishing effort data could be made in both snow crab and red king crab assessments. There may well be problems in defining meaningful effort and CPUE indices, particularly for the by-catch fleets, but there are potentially great benefits in doing so. Firstly, it may be important to understand spatial processes in the stocks and the fisheries. It is probably unrealistic to expect that spatial processes could be incorporated in the LBA models in the near future, but statistical analyses of CPUE and effort data may well be informative about shifts in the location of the stock or the fishery, which in turn may be informative about selectivity and mortality processes that are addressed within the models. Secondly, it is important to gain a better insight into the contribution of by-catch mortality to overall mortality in BSAI crab stocks. In this case direct use of effort data in the LBA may be possible. According to Dew & McConnaughey (2005), by-catch mortality of females could have been an important factor in the decline of Bristol Bay red king crab stocks after 1980. Whilst this is not the only view of causes of decline, it does highlight the importance of understanding by-catch mortality. Dew & McConnaughey (2005) also question the representation of 'red bag' catches in estimates of the by-catch component of red king crab removals. Again, whether or not this proves to be a real source of bias in the data, it highlights the importance of understanding and quantifying the contribution of by-catch to overall fishing mortality. This is important for both assessments and OFDs.

ToR (b): Recommendations of improvements to proposed overfishing definitions

Recommendations about how the proposed OFD could be developed or improved are scattered through the preceding sections. The main points are highlighted again below, together with some additional recommendations on the simulations and OFD framework.

- The proposed OFD can be accepted as a framework, but it is urgently necessary to make progress on defining values for the reference points (F_{MSY} or proxies) and defaults for parameters α, β and γ. This will only be possible when a satisfactory interim definition for ESB is derived and the proposed OFD is tested in conjunction with realistic harvest strategies. Retention of the existing OFD is not an option.
- Protocols are needed for dealing with the addition of new annual assessment estimates to existing OFDs. This might involve use of running averages or a cycle of update assessments and full revisions to the OFD.
- In the short-term, there needs to be a prescriptive interim definition of how ESB is to be calculated and used in the OFD and simulations. Mature male biomass is a strong candidate for this definition. Subsequent improvements to the definition of ESB, *e.g.* based on the analysis of survey data or the outcome of an IBM, should be extensively reviewed and documented before being adopted in a revised OFD.

- In the immediate term, simulations to test the performance of the proposed OFD and to determine appropriate values for α, β and γ in the MSY control rule must include realistic harvest strategies. To do otherwise is either to take the view that no precautionary buffer is needed between F_{OFL} and the target F or to determine OFDs that allow the State very little room for manoeuvre in setting harvest strategies to meet precautionary and other objectives.
- There needs to be agreement on the criteria used to test OFDs in the simulations. Rebuilding time for a depleted stock would be an important criterion, but trade-off statistics involving the level and variability of yield in the short- and long-term are also needed.
- The default OFD of Restrepo *et al.* (1998) should be included in comparisons of the performance of different OFDs. This is less precautionary than the proposed OFD framework, but specifies a higher value of MSST given M<0.5 (see Figure 1). It will need to be established that the increased complexity involved in the proposed OFD (*i.e.* the α and β parameters) offers significant improvements over the default OFD in terms of the agreed test criteria.
- The simulations should include a range of starting values for stock biomass, in terms of fractions of B_{MSY}. This is because rebuilding is not the only function of an effective OFD – it is also intended to define sustainable exploitation of a healthy stock.
- There needs to be agreement between ADF&G and NMFS teams on the interpretation of the available evidence on biological processes (*e.g.* moulting/mating cycles) in BSAI crab stocks. In cases of genuine uncertainty, this should be included in sensitivity analyses. Comparisons between the ADF&G and NMFS simulation modelling approaches would be facilitated by the use of common starting points for the simulations (fractions of B_{MSY}).
- Simulation testing of the performance of proposed OFDs must take appropriate account of observation error, *i.e.* the difference between the simulated 'reality' and the 'observations' used in applying the OFD. Ideally, this would involve simulating the survey and assessment processes (see Figure 2), but simulating the errors directly would probably be adequate given knowledge of their likely magnitude and autocorrelations.
- For Tier 5 stocks, it is not appropriate to use the average catch for a single fixed period of years to define the OFL. The period should be defined separately for each stock based, where possible, on four criteria: (i) stability of catches; (ii) lack of trend in CPUE; (iii) lack of trend in fishing effort; and (iv) a stable spatial distribution of the fishery, showing no expansion or shift in the distribution of fishing effort. If possible, the use of these criteria should be supported by simulation studies.
- In the medium-term, consideration should be given to including uncertainty measures within the Tiers of the OFD. This might involve estimating probabilities for the current location of a stock in relation to status determination criteria rather than just using point estimates.
- In the medium-term, consideration should be given to reducing the complexity of the OFD. For example, would a flat F_{MSY} control rule (in conjunction with a

precautionary harvest strategy) perform as well as the proposed MSY control rule, thus removing the need to define α and β?

- In the medium-term, further attention should be paid to the role of by-catch mortality in determining the effectiveness of the OFD. At present, by-catch mortality is seen as a context for the OFD, but it is a context that readily changes in response to the fortunes of other fisheries. Ideally, by-catch mortality should be included in the F that is compared with F_{OFL} (with implications for defining F_{OFL}). This has the disadvantage that the by-catch F will need to be projected before the TAC can be calculated, but it does allow an OFD that does not require revision each time the by-catch mortality regime changes.
- In the long-term, it is recommended that OFDs include multi-species considerations. This is necessary because: (i) by-catch of BSAI crabs in trawl fisheries, whether or not removed in the form of commercial landings, is potentially an important contribution to overall fishing mortality; and (ii) the catches of even directed fisheries are often mixed in species composition. Construction of robust single species OFDs is a necessary precursor to more complex management systems, but management of trade-offs between competing objectives for mixed or interacting fisheries potentially offers the capacity to maximize overall conservation (and revenue) benefits.
- Another long-term objective should be to include spatial considerations within the OFD. For spatially structured stocks, the consequences of fishing mortality for future recruitment depend heavily on when and where the mortality occurs. As an example, the hypothesis of Dew & McConnaughey (2005) that trawling in the south-eastern Bering Sea has disrupted the ‘endless belt’ reproductive strategy of red king crabs holds strong implications for spatial management. In this case the implications apply to the by-catch fleets, but it is easy to envisage cases where spatial management considerations would apply to the directed fishery. The inclusion of spatial management criteria in an OFD implies that the current type of MSY control rule would no longer be appropriate. Equilibrium recruitment at a given level of fishing mortality is the determinant of F_{MSY}, but in the case of spatial management there is no single F_{MSY}, and recruitment is determined by considerations beyond a ‘global’ SRR.

ToR (c): Review of model configurations, formulations and methods used to account for uncertainty

The model configurations, formulations and methods used to account for uncertainty have already been reviewed under ToR (a) alongside the OFDs and simulation models. Model structures (LBAs and projection models) appear to be sufficient and appropriate to the life-histories of BSAI crab stocks, although there remain some disagreements about the details of some biological processes. Likewise, model fitting procedures for the LBAs appear satisfactory, although this was not an aspect that could be examined in detail during the review. The review team did not take up the offer of access to AD Model Builder and Fortran code for the models. This was partly because there would have been insufficient time to read and thoroughly understand the code, but also because the presentations and documentation for the review made it clear that the technical expertise of the modelling teams was not in question.

ToR (d): Review of input parameters used in simulation models

Input parameters used in the simulation models included biological and fishery parameters and SRRs. Again, these have largely been discussed already under ToR (a). It is worth re-iterating that there remain several important points of difference between ADF&G and NMFS teams in interpretation of the scientific evidence on biological processes in BSAI crab stocks, and it is important that these differences are resolved. For progress in parameterising the OFDs it is necessary to draw up a clear and unequivocal agreed framework of crab life-history processes. Irresolvable points of genuine uncertainty should (a) be accounted for in sensitivity analyses, and (b) serve as a focus for future research efforts.

Opportunity to examine important fundamental biological parameters was limited during the review. In common with many if not most exploited species, M is poorly known for BSAI crab stocks, and there is a slender basis for drawing inferences about likely values. Values of 0.2 or 0.3 used as proxies for F_{MSY} for Tier 4 stocks and alternative values for M used in the simulation models appear to have been derived from considerations of longevity. The estimates are satisfactory to the extent that they are at least plausible. There is no obvious basis for their revision, although it should be noted that there should be consistency of selected best values across assessment models, simulation models and OFDs, unless precautionary considerations dictate otherwise.

Growth parameters are also poorly known, which could have important implications for assessments, simulation models and estimation of $F_{x\%}$ values in the selection of F_{MSY} proxies. The sensitivity of the models to assumptions about growth are certainly worth exploring, although it is also possible that internal consistency in the assessment-simulation-OFD model complex may be more important than absolute lack of bias when determining the most effective OFD.

ToR (e): Suggested research priorities

Some priorities for research have already been identified in the preceding sections. These are collected together below, together with some further suggestions aimed at improving the understanding of essential population and fishery dynamics necessary to formulate best management practices.

- The first and most urgent research priority is to determine the appropriate functional form for calculating a measure of ESB that is proportional to TFEP. This is required for the MSY control rules and for the SRRs that are used to derive F_{MSY} values and proxies, to find likely default values to parameterise the OFDs and to test the performance of OFDs. As described under ToR (a), mature male biomass may be an appropriate candidate for an interim definition of ESB, but cross-correlations between the various options (*e.g.* B_0 mating ratio applied to mature female biomass) should be examined to determine the extent to which they are measuring the same dimension of variability (*i.e.* measures that go up and down in rough concert with various alternative formulations of spawning potential, even if the correspondences are non-linear). In the medium- to long-term, construction of an IBM and analysis of survey data are the most likely routes to improving the formulation of ESB (see under ToR (a) above).

- Field estimation of clutch fertilisation rates during annual surveys is important for improving the understanding of the determinants for TFEP. This would need to be carried out over a number of years (preferably contrasting) for meaningful conclusions to be drawn. The study may shed light on whether it is possible to use records of egg colour to drawn conclusions about variations in fertilisation rates over a longer series of years.
- Depletion experiments using the survey vessel and gear should be used to obtain estimates of survey efficiency and selectivity. This information could be incorporated into the assessments for snow crab and red king crab, allowing more robust and parsimonious models. As pointed out by Nick Caputi during the review meeting, catch rate data from intensively fished areas may also allow estimation of selectivity and catchability parameters for commercial fishing operations.
- It is recommended that spatial management considerations be included in the future development of OFDs (see recommendation under ToR (b)).
- It is recommended that multi-species considerations be included in the future development of OFDs (see recommendation under ToR (b)).
- Further use of CPUE and fishing effort data could be made in the assessments. It is recommended that effort and CPUE data be collated for both directed fisheries and by-catch fleets and that trends in these data be examined in a spatial context. Generalised linear modelling or other appropriate statistical techniques could be used to extract annual and spatial signals. Use of these signals in the assessments should be investigated, *e.g.* to improve understanding of the contribution of by-catch mortality to overall levels of fishing mortality.
- It is recommended that there be an investigation of the sensitivity of biological reference points and the performance of the OFD to uncertainty about biological parameters, especially growth.
- If lack of information about growth is determined to be an important source of uncertainty about the effectiveness of assessment and precautionary fishery management, it is recommended that field tagging studies be used to estimate growth increments and moulting frequency in BSAI crab stocks. Properly designed tagging studies, *e.g.* involving sufficiently large samples of crabs tagged in relatively shallow waters, could also shed light on rates of natural mortality. Lipofuscin measurements could also be used to investigate the relationship of size with age, although this may depend on the development of routine methods of lipofuscin determination that can be applied to large samples.
- Under ToRs (a) and (b) it was recommended that that the ADF&G and NMFS teams need to reach agreement on the interpretation of the available information on some key

biological processes in BSAI crabs, *e.g.* the interval between moulting and participation in mating by new shell male snow crabs. Where genuine uncertainty remains about key biological processes, this should be used as a focus for new research on crab life-histories, particularly if simulations reveal that the OFD is sensitive to this uncertainty.

- Further research is needed into the relative roles played by fishing operations and changes in the climatic and oceanographic regime in determining past trends in BSAI crab stocks. A consensus on this contentious issue is needed before progress can be made in determining agreed protocols for detecting the effects of regime shifts on crab productivity and in determining the appropriate management response to such shifts.

Conclusions

A great deal of effective research and analytical work has been undertaken to put forward proposals for an OFD that is a vast improvement on the current flawed OFD. In my view this effort has been largely successful in that the proposed OFD is one that can now be accepted as a framework. The challenge now is to progress the parameterisation of the framework to the point where it can be implemented. This means finding default values for the parameters within each tier, *e.g.* stock-specific values for α and β in Tier 3 and group-specific values of α, β and γ in Tier 4, and proxy values for F_{MSY} in Tier 3. At present it is not yet possible to recommend particular values for these parameters, because certain issues need to be resolved before the performance of the OFD under any given parameterisation can effectively be tested. The approach to finding appropriate defaults using simulation models can nevertheless be endorsed as sound in principle and correctly specified in practice (in terms of model structures).

There are two main obstacles to finding default values for the OFD parameters. The first is that there is not yet a satisfactory measure of ESB to used in the MSY control rules and in the SRRs within the simulations. Ideally, ESB should be a measure that is proportional to TFEP; none of the candidate measures considered so far meet this criterion. Recommendations are made in this report for a simple interim measure that could be used immediately and for research aimed at finding more satisfactory long-term solutions.

The second main obstacle is that the role of the harvest strategy in determining the performance of the OFD has not yet been considered. Simulations have so far treated MSY control rules as if they were harvest strategies. On the one hand, this does not recognise the requirement for the State to maintain a precautionary buffer between F_{OFL} and the target F. On the other hand, choice of an MSY control rule on this basis is likely to result in an OFD that places undue constraints on the capacity of the State to manage BSAI crab fisheries according to precautionary and other objectives. What is needed is for MSY control rules to be tested in conjunction with realistic State harvest strategies. Only then can an OFD be selected that serves its proper role of providing limits rather than targets for safe management.

From this I conclude that there remains some work to be done before the OFD can be accepted for implementation. However, the obstacles are not insurmountable, even in the short-term. Given the urgent need for the current OFD to be replaced by a more satisfactory alternative, the Crab Plan Team and its Work Group should be encouraged to select an interim measure for ESB and to use simulations of MSY control rules in conjunction with harvest strategies to select appropriate parameters for an OFD that can be implemented in the short-term. This report also contains recommendations for improvements and developments to the OFD in the medium- to long-term and for supporting research, but these should not be seen as reasons to delay implementation.

Acknowledgements

I would like to thank Anne Hollowed, Jim Ianelli, Lou Rugolo, Diana Stram, Shareef Siddeek, Jack Turnock and Jie Zheng for effective, interesting and good humoured presentations and discussions during the review meeting. Aric Bickel and Manoj Shivlani, University of Miami CIE, steered me and my fellow reviewers through the complications and pitfalls of attending the review and subsequent meetings. I would also like to thank my fellow reviewers, Nick Caputi and Patrick Cordue, for stimulating discussions during the review meeting and for providing notes for my presentations at the CPT and SSC meetings.

References

Clark, W.G., 1991. Groundfish exploitation rates based on life history parameters. *Canadian Journal of Fisheries & Aquatic Sciences*, **48**, 734-750.

Dew, C.B. & McConnaughey, R.A., 2005. Did trawling on the brood stock contribute to the collapse of Alaska's king crab? *Ecological Applications*, **15**, 919-941.

Maunder, M.N., 2003. *Review of the stock assessment and harvest strategy for eastern Bering Sea snow crab.* CIE, University of Miami.

NPMFC, 2006. *Workshop Report: Crab Overfishing Definitions Inter-agency Workshop. February 28-March 1, 2006, Alaska Fisheries Science Center, Seattle, WA.* NPMFC, Anchorage.

Restrepo, V.R., Thompson, G.G., Mace, P.M., Gabriel, W.L., Low, L.L., MacCall, A.D., Methot, R.D., Powers, J.E., Taylor, B.L., Wade, P.R. & Witzig, J.F., 1998. *Technical guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act.* NOAA Technical Memorandum NMFS-F/SPO-##.

Rugolo, L., 2004. *North Pacific Fisheries Management Council Bering Sea/Aleutian Islands King and Tanner Crab Working Group: Draft Statement of Work*. NMFS/ADF&G, Kodiak/Seattle/Juneau.

Rugolo, L. 2006. *Statement of Work: NPFMC BSAI King and Tanner Crab Working Group*. www.afsc.noaa.gov/refm/docs/2006/crab/Statement%20of%20Work.ppt

Siddeek, M.S.M. & Zheng, J., 2006. *Reference point estimation analysis for the Bering Sea and Aleutian Islands (king and Tanner) crab revised fisheries management plan.* ADF&G, Juneau.

Turnock, B.J. & Rugolo, L.J., 2005. *Stock assessment of eastern Bering Sea snow crab.* NMFS, Seattle/Kodiak.

Turnock, B.J. & Rugolo, L.J., 2006a. *Analysis of proposed overfishing tier system for BSAI king and Tanner crab stocks.* NMFS, Seattle/Kodiak.

Turnock, B.J. & Rugolo, L.J., 2006b. *Unresolved issues concerning proposed overfishing definitions for Bering Sea and Aleutian Islands king and Tanner crab stocks: National Marine Fisheries Service.* NMFS, Seattle/Kodiak.

Zheng, J., 2004. *Bristol Bay red king crab stock assessment in 2004.* ADF&G, Juneau.

Zheng, J., 2006. *Issues dividing the Crab Work Group.* ADF&G, Juneau.

Zheng, J. & Kruse, G.H., 2006. Recruitment variation of eastern Bering Sea crabs: Climate-forcing or top-down effects? *Progress in Oceanography*, **68**, 184-204.

APPENDIX 1: Bibliography of materials provided during the review meeting

The key documents referred to during the review are listed below:

Dew, C.B. & McConnaughey, R.A., 2005. Did trawling on the brood stock contribute to the collapse of Alaska's king crab? *Ecological Applications*, **15**, 919-941.

Maunder, M.N., 2003. *Review of the stock assessment and harvest strategy for eastern Bering Sea snow crab.* CIE, University of Miami.

NPMFC, 2006. *Workshop Report: Crab Overfishing Definitions Inter-agency Workshop. February 28-March 1, 2006, Alaska Fisheries Science Center, Seattle, WA.* NPMFC, Anchorage.

Restrepo, V.R., Thompson, G.G., Mace, P.M., Gabriel, W.L., Low, L.L., MacCall, A.D., Methot, R.D., Powers, J.E., Taylor, B.L., Wade, P.R. & Witzig, J.F., 1998. *Technical guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act.* NOAA Technical Memorandum NMFS-F/SPO-##.

Rugolo, L., 2004. *North Pacific Fisheries Management Council Bering Sea/Aleutian Islands King and Tanner Crab Working Group: Draft Statement of Work.* NMFS/ADF&G, Kodiak/Seattle/Juneau.

Rugolo, L. 2006. *Statement of Work: NPFMC BSAI King and Tanner Crab Working Group*. www.afsc.noaa.gov/refm/docs/2006/crab/Statement%20of%20Work.ppt

Siddeek, M.S.M. & Zheng, J., 2006. *Reference point estimation analysis for the Bering Sea and Aleutian Islands (king and Tanner) crab revised fisheries management plan.* ADF&G, Juneau.

Turnock, B.J. & Rugolo, L.J., 2005. *Stock assessment of eastern Bering Sea snow crab.* NMFS, Seattle/Kodiak.

Turnock, B.J. & Rugolo, L.J., 2006a. *Analysis of proposed overfishing tier system for BSAI king and Tanner crab stocks.* NMFS, Seattle/Kodiak.

Turnock, B.J. & Rugolo, L.J., 2006b. *Unresolved issues concerning proposed overfishing definitions for Bering Sea and Aleutian Islands king and Tanner crab stocks: National Marine Fisheries Service.* NMFS, Seattle/Kodiak.

Zheng, J., 2004. *Bristol Bay red king crab stock assessment in 2004.* ADF&G, Juneau.

Zheng, J., 2006. *Issues dividing the Crab Work Group.* ADF&G, Juneau.

Zheng, J. & Kruse, G.H., 2006. Recruitment variation of eastern Bering Sea crabs: Climate-forcing or top-down effects? *Progress in Oceanography*, **68**, 184-204.

Further documentation available to the reviewers, including presentations given to the crab overfishing workshop is given at:

http://www.afsc.noaa.gov/refm/stocks/CrabWs.htm

APPENDIX 2: Agenda for the review meeting

Center of Independent Experts
Alaska Crab Overfishing Definitions

April 24 - 29, 2006

Alaska Fisheries Science Center, Seattle, WA

Apr 14th 2006 Draft Agenda

Purpose: To solicit expert advice on proposed overfishing definitions for Bering Sea and Aleutian Islands crab stocks. We are requesting a review of issues critical to formulating new overfishing definitions, biological reference points, input parameters, modeling approaches and methods to deal with uncertainty.

DAY 1 (Center Director's Conference Room)

8:00 Coffee and informal discussions

8:30 Introductions - Charge for the CIE –Hollowed

8:50 History of crab management - current overfishing definitions and need for revision - Stram or Designee

9:10 Overview of proposed revisions - Working group

- Working group Statement of Work (20 min) - Rugolo
- Tier System review (20 min) - Zheng
- Brief Description of Snow Crab Assessment (40 min) -Turnock

10:30 Break

10:30 – 12:00 Overview continued – working group

- Brief Description of Red King Crab Assessment (40 min) -Zheng
- Projection Model structure (Siddeek and / or Turnock)

12:00 – 1:00 Break for lunch

1:00-1:30 Overview continued – working group

- Approaches to estimate proxy values for biological reference points – Turnock
- Approaches to estimate proxy values for biological reference points - Siddeek

1:30 – 2:00 Review Workshop Report and Recommendations on crab biology – Stram or designee

2:00 – 2:30 Review of Workshop Report and Recommendations on crab modeling - Ianelli

2:30 Break

2:45-3:45 Review of information available for managed crab stocks - Rugolo

3:45 – 5:00 Performance of Tier System Preliminary results

- Red King Crab – Siddeek
- Red King Crab – Turnock

DAY 2 (CD Conference Room)

8:30 Coffee and informal discussions

8:30 – 10:00 Performance of Tier System Preliminary results continued

- Snow Crab – Turnock
- Snow Crab – Siddeek
- Blue King Crab/Golden Crab - Siddeek

10:00 Break

10:30 – 12:00 Questions and Answers for panel.

12:00 Lunch

1:00 – 5:00 Open question and answer session – or independent work sessions with CIE reviewers.

DAY 3 (CD Conference Room)

8:30 Coffee and informal discussions

9:00 Open question and answer session – or independent work sessions with CIE reviewers.

DAY 4 (CD Conference Room)

8:30 Panel discussions and writing team – NMFS and ADF&G biologists return to offices but remain on cal questionsl

DAY 5 (CD Conference Room)

8:30 Panel discussions and writing team – NMFS and ADF&G biologists return to offices but remain on call to answer questions

APPENDIX 3: Statement of Work

STATEMENT OF WORK

April 19, 2006

General

The Alaska Fisheries Science Center (AFSC) requests review of proposed overfishing definitions and simulation models used to evaluate biological reference points for Bering Sea and Aleutian Islands King and Tanner crab stocks. The North Pacific Fishery Management Council (NPFMC) has determined that the existing overfishing definitions for Bering Sea and Aleutian Islands King and Tanner crab stocks need revision. The AFSC is seeking review of the population dynamics models developed for revising the overfishing definitions.

There are currently 22 Bering Sea and Aleutian Islands crab stocks under the Federal Bering Sea Aleutian Island Crab Fishery Management Plan (FMP) of which 7 are considered major stocks. Four of the seven major crab stocks have been declared overfished and rebuilding plans developed within the last 7 years. Of the remaining three stocks, only one has been relatively stable at a low level, another has maintained stable catch for several years, however, even for this stock it appears recruitment may be declining. While the remaining stock has increased, survey abundance estimates have low precision and the fishery is closed due to bycatch concerns. There is no consensus on the principal cause of declines in Bering Sea crab stocks.

A panel of 3 consultants is requested for this review. The panel will need to be thoroughly familiar with various subject areas involved in analytical stock assessment, including population dynamics theory, length based stock assessment models, rebuilding analyses, estimation of biological reference points and harvest strategy modeling for invertebrates, as well as invertebrate biology. The CIE consultants will travel to Seattle, Washington to meet with the four member Interagency Work Group charged with developing the new overfishing definitions. We request that one member of the Panel should be present at the May meeting of the NPFMC Crab Plan Team in Anchorage, Alaska. The report generated by the consultants should include:

a. A statement of the strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.
b. Recommendations for improvements to proposed overfishing definitions or alternative definitions,
c. A review of the model configurations, formulations and methods used to account for uncertainty.
d. A review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in simulation models.

e. Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

AFSC will provide copies of the NPFMC Work Group statement of work, proposed overfishing definitions, preliminary results of simulations, discussion of input parameters, a copy of the code for the snow crab stock assessment, and the AD Model Builder and Fortran code used for reference point estimation. The panel will meet with scientists from the Alaska Fisheries Science Center and the Alaska Department of Fish and Game from April 24 to April 28, 2006, in Seattle, Washington.

Expected Products:

- One member of the panel will attend the May meeting of the Crab Plan Team to discuss the panels findings regarding the strengths and weaknesses of proposed definitions and modeling approaches.
- No later than June 1, 2006, panelists will submit a written report of findings, analysis, and conclusions. The report should be addressed to the "UM Independent System for Peer Reviews", and sent to David Die, UM/RSMAS, 4600 Rickenbacker Causeway, Miami, FL 33149 (or via email to ddie@rsmas.miami.edu).

Signed________________________

Date____________

ANNEX I: REPORT GENERATION AND PROCEDURAL ITEMS

1. The report should be prefaced with an executive summary of findings and/or recommendations.

2. The main body of the report should consist of a background, description of review activities, summary of findings, and conclusions/recommendations.

3. The report should also include as separate appendices the bibliography of materials provided by the Center for Independent Experts and the center and a copy of the statement of work.

4. Individuals shall be provided with an electronic version of a bibliography of background materials sent to all reviewers. Other material provided directly by the center must be added to the bibliography that can be returned as an appendix to the final report.

Please refer to the following website for additional information on report generation: http://www.rsmas.miami.edu/groups/cimas/Report_Standard_Format.html

Alaskan Crab Overfishing Definitions Review

Seattle, Washington

24-28 April 2006

Dr Nick Caputi

Department of Fisheries (Western Australia)
Western Australian Fisheries and Marine Research Laboratories
PO Box 20, North Beach, WA 6920, Australia

Representing the Center of Independent Experts, University of Miami

Executive Summary

The Alaska Fisheries Science Center (AFSC) requested a review of proposed overfishing definitions and simulation models used to evaluate biological reference points for Bering Sea and Aleutian Islands King and Tanner crab stocks. The North Pacific Fishery Management Council (NPFMC) has determined that the existing overfishing definitions for Bering Sea and Aleutian Islands King and Tanner crab stocks needed revision. The AFSC sought a review of the population dynamics models developed for revising the overfishing definitions.

There are currently 22 Bering Sea and Aleutian Islands crab stocks under the Federal Bering Sea Aleutian Island Crab Fishery Management Plan (FMP) of which 7 are considered major stocks. Four of the seven major crab stocks have been declared overfished and rebuilding plans developed within the last 7 years. Of the remaining three stocks, only one has been relatively stable at a low level, another has maintained stable catches for several years, however, even for this stock it appears recruitment may be declining. While the remaining stock has increased, survey abundance estimates have low precision and the fishery is closed due to bycatch concerns. There is no consensus on the principal cause of declines in Bering Sea crab stocks.

A panel of three consultants undertook the review. The panel met with scientists from the Alaska Fisheries Science Center and the Alaska Department of Fish and Game charged with developing the new overfishing definitions from April 24 to 28, 2006, in Seattle, Washington. The crab team presented the key aspects of their research on the first three days. Throughout the presentations the CIE panel asked detailed questions on issues of the stock assessment related research that was presented. All members of the crab team answered questions and expanded on some aspects of the stock assessment.

AFSC provided access to a number of relevant papers that were listed on their web site www.afsc.noaa.gov/refm/stocks/CrabWs.htm and provided some additional documents by email. The key papers that focused on the area of review were:

- Statement of work for working group.
- Description of proposed overfishing definition tier system.
- Stock assessments for Red King Crab and Snow Crab.
- Working group position papers.
- Workshop report recommendations.
- Projection model results.

This CIE review team was asked to focus on:

a. A statement of the strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.
b. Recommendations for improvements to proposed overfishing definitions or alternative definitions,

c. A review of the model configurations, formulations and methods used to account for uncertainty.
d. A review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in simulation models.
e. Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

Federal legislation requires an overfishing definition (OFD) that specifies whether the stock is overfished and whether there is overfishing occurring. The proposed system represents a significant improvement as it is based on the current NPFMC groundfish system which has been reviewed and tested. A buffer is incorporated between the overfishing limit (FOFL) and the target F level as required on the National Standard guidelines 1 (NSG1). In the current crab tier system there is no buffer between the target F and FOFL.

The proposed framework is comprehensive having five tiers which take into account the level of knowledge and uncertainty about the stocks being managed. However the uncertainty within a tier has not been thoroughly taken into account and should be considered when considering the overfishing and overfished definitions and the strategies for rebuilding. For Tiers 1 to 4 there are three levels of stock status with a corresponding target fishing mortality rate corresponding to the overfishing limit (FOFL).

The annual assessment of the stock provides for an annual revised estimate of the OFD levels with a revision of the model approach, the parameters of the model and the new year's data. This provides the 'best' indication of the status of stock. However this could also be viewed as a weakness of the proposed OFD approach in that the OFD can change with each year's stock assessment. A two-stage approach should be considered for each year's stock assessment: (1) a comparison of the latest year's stock level and exploitation with the OFD level set in the previous year's definition for overfished and overfishing; and (2) undertake a revised stock assessment which may include a new model approach, revised biological parameters as well as the addition of the usual new year's data.

Modelling of the proposed overfishing tier system by the two modeling groups is viewed as a strength in the process of determining the OFD in that it provides a comparison of alternative approaches, different set of assumptions about the features in the model such as the measure of stock (B) which is the basis of the overfished assessment and the type of the stock-recruitment relationship (SRR). However to gain maximum benefit from the two modeling approaches it is important to undertake critical analysis of the results and provide a revision and improvement to the models. Some revision of the models has occurred but no consensus on the optimum model has been reached.

The projection model to compare rebuilding strategies and different parameters should have the same starting biomass for each simulation. This was undertaken by Turnock and Rugolo (2006) but Siddeek and Zheng (2006) use a different starting value (beta x Bmsy) for some of the different comparison of parameters. This means that some of the simulations are not comparable in assessing the parameters. The different levels of alpha

(0 to 0.1) tested show little difference in rebuilding time and long-term mean yield so any value in this range appears satisfactory. One of the weaknesses in the new OFD approach in the choice of alpha and beta in the OFD are somewhat arbitrary and default levels of 0 and 0.2 can be used in the absence of evidence to indicate that there are more appropriate measures.

The projection model tests the harvest rule from the proposed Tier system as well as the current OFL and the current ADFG harvest strategies. The simulation confirms that the current OFL is not sustainable and there is a good comparison of a large number of rebuilding strategies.

As you move down the Tiers 2 to 4, the models are more sensitive to scientist decisions as less information is available and hence require additional simulations to assess the relative merits of the model. Tier 5 should consider effort data in setting a target catch level. For example, has there been an increase or decrease in effort for the periods under consideration for setting the target catch? If there is considerable annual variation in recruitment then this increases the chance of overfishing if there is a series of below-average recruitment. Simulation analyses associated with this Tier should be conducted to assist in determining a sustainable control rule.

Some additional recommendation to assess the OFDs:

- An assessment should be made of the short-term impact of rebuilding on catch compared to the rebuilding time.
- There is a need to consider variability in the parameters, observation error, and hence the uncertainty associated with the current status relative to the decision rules within each of the tiers and the uncertainty associated with rebuilding strategies so that managers can be aware of the variability associated with these assessments.
- Additional simulations are required to assess the relative merits of the OFD models as you move down the tiers 2 to 4, the models are more sensitive to scientist decisions as less information is available. Tier 4 requires additional simulations to assess an additional parameter (gamma).

The measurement of egg production is particularly difficult for the Alaskan crab fishery which is a male only fishery resulting in a large numbers of mature females that are unmated, females with clutches that are not filled, females with unfertilized eggs, and barren and senescent females. These are all indicators of a relatively much lower abundance of mature males compared to mature females which results in the mature males being the limiting factor in the determining the egg production. Hence the annual mature male abundance (taking into account sperm variation with size) in the appropriate location may be the key determinant to egg production and should be considered as a possible indicator of egg production. The indicator used by Turnock and Rugolo (2006) take into account the fact that mature males are limited in determining effective mature female biomass but then it adds the effective male mature biomass which does not appear appropriate.

The cause of the reduction in the king crab stocks since the 1980's is critical in determining what are the target Bmsy levels. If the reduction is due to a regime shift then basing the Bmsy on the lower levels of mature biomass since the 1980's is appropriate. There is evidence of the negative effects of the increase in trawling since 1980, particularly in the most productive spawning grounds off Unimak and Amak Islands, on the breeding stock. However it may not be possible to restrict trawling from the more productive spawning areas in which case basing the Bmsy on the lower levels of mature biomass since the 1980's is still appropriate as the breeding stock will not return to the levels of the 1970's.

An adaptive management approach should be considered to assess the effects of fishing on these productive grounds by closing an appropriately-sized area to trawling to determine the impact on the stock in that area. The two competing hypotheses on decline of the king crab stocks since the 1980's, i.e. regime shift and the effects of increased targeted and trawling, may both be contributing to the decline in recruitment. Many stocks quite often collapse when there is the combined effect of poor environmental conditions at a time when the breeding stock is reduced due to changes in fishing practices.

The SRR is also affected by the years chosen to assess the fit and the significant change to the recruitment pattern before and after 1976. Irrespective of whether this change is due to a regime shift or the effects of trawling, there will be a change in the shape of the SRR and this should be taken into account.

The choice of the stock-recruitment relationship (SRR) is important in the stock assessment of the Alaskan crab fisheries and both modeling groups have given this issue a significant level of attention. The Maximin Clark (1991) method provides a basis to assess different steepness levels of the SRRs when there is no empirical data available. However in many cases there are some data available to at least make a choice about whether the SRR is likely to be a Ricker or Beverton-Holt curve.

As the relative size of mature males and females is import in the mating process, it is important to monitor the changes in mean size and length frequency for mature males and females that occur. The ratio of mature male to mature female mean size could also be used to measure the relative changes in mean size.

The Turnock and Rugolo (2006) population models have a large number of parameters estimated and it appears these could be significantly reduced eg there appears to be little biological basis for having separate male and female recruitment indices (even if they 'were constrained to be similar'). The annual recruitment of males and females should be similar and set at appropriate sex ratio if the recruitment sex ratio is not 1:1. Also the biological basis for having different selectivities for new and old shell is not clear. Annual parameters are estimated for selectivities and again it is not clear why selectivity should change every year. The use of different natural mortality levels for 3 different periods for males and 4 different periods for females does not appear to be biologically sensible (Zheng 2006).

Estimation of survey catchability for snow crabs using underbag have been undertaken. However this may not provide a complete assessment of the catchability. The use of a depletion experiment should be considered to estimate survey catchability for different sizes, shell condition and sexes. Environmental factors can have a significant impact on the efficiency of the gear and it would be useful to have an assessment of this issue. The key environmental indices during the surveys should be summarized so that the potential biases in the indices are identified.

Some suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices include:

1. As mature males may be the limiting factor in the determining the egg production, the annual variation in the mature male abundance should be considered in modelling as a possible indicator of egg production.
2. Depletion experiments should be considered to estimate survey catchability for different sizes, shell condition and sexes.
3. A depletion analysis of some blocks that are heavily fished during a season such that there is a significant decline in catch rate due to the effects of fishing could provide some valuable insights into some fishery dynamics. A comparison of the daily retained male CPUE in a block (or groups of blocks) and the cumulative legal catch removed from that block over the period that the fishery operates enables an estimate of the residual legal biomass at the end of fishing, the catchability of the male crabs and the exploitation rate.
4. A depletion analysis may also be applied to assess the impact of fishing on discards if there is sufficient observer data on the daily catch rate of discards in a heavily fished block(s) and an estimate of discard numbers can be made from those block(s). A significant decline in the discard rate during the course of fishing would indicate a significant level of discard mortality.
5. The change in management of the fishery to an individual transferable quota (ITQ) is likely to result in high grading and hence an increase in the rate of discarding and hence associated discard mortalities. Consideration should also be given to retaining some of the discards by providing a separate quota for discards. If there is a high mortality (50-100%) associated with discards it may be worth retaining some of them (if there a market for them) and reducing the ITQ for the first-grade crabs. This issue is also related to Recommendation 7.
6. While considerable research on escape gaps and subsequent changes have been undertaken on escape gaps, it appears that there is still considerable retention of undersize crabs, most (50-100%) of which may die as a result of being captured. This makes it imperative to undertake further research (if necessary) to choose the number and size of the escape gaps that maximizes the escape of undersize male and female crabs even if it means that some of the smaller legal-size males are allowed to escape. Additional research on the

handling practices (dropping crabs on a hard surface from a height of greater than 4 ft) onboard should also be undertaken to assess if there are ways to improve handling practices to increase survival of discards.

7. An evaluation should be undertaken on the merits of retaining some female king crabs that are marketable as part of the catch. There appears to be a surplus number of mature females relative to the number of mature males in the fishery resulting in unmated and senescent females. These females could contribute to significant loss of productivity due to density dependent mortality and growth, particularly if habitat is limiting. A modeling of harvest strategies should be examined that includes the retention of an appropriate quantity of females that results in an optimum ratio of mature males to mature females and hence a reduction in unmated mature females.
8. The modeling of the shell condition is a critical part of the population dynamics of the crab fishery as it affects the catch that is targeted and retained, molting, growth, maturity and the mating dynamics. There appears to be uncertainty about the relationship that has been assumed between shell condition and time since last moulting and this relationship needs to be examined further.
9. An economic assessment of the fishery should be undertaken in conjunction with the stock assessment modelling to assess ways to improve the economic performance of the fishery. The maximum economic yield (MEY) which is less than MSY should be considered as a performance indicator for the fishery as it would be a more conservative indicator.

Background

The Alaska Fisheries Science Center (AFSC) requested a review of proposed overfishing definitions and simulation models used to evaluate biological reference points for Bering Sea and Aleutian Islands King and Tanner crab stocks. The North Pacific Fishery Management Council (NPFMC) has determined that the existing overfishing definitions for Bering Sea and Aleutian Islands King and Tanner crab stocks needed revision. The AFSC sought a review of the population dynamics models developed for revising the overfishing definitions.

There are currently 22 Bering Sea and Aleutian Islands crab stocks under the Federal Bering Sea Aleutian Island Crab Fishery Management Plan (FMP) of which 7 are considered major stocks. Four of the seven major crab stocks have been declared overfished and rebuilding plans developed within the last 7 years. Of the remaining three stocks, only one has been relatively stable at a low level, another has maintained stable catch for several years, however, even for this stock it appears recruitment may be declining. While the remaining stock has increased, survey abundance estimates have low precision and the fishery is closed due to bycatch concerns. There is no consensus on the principal cause of declines in Bering Sea crab stocks.

A panel of three consultants was requested for this review. The panel was familiar with various subject areas involved in analytical stock assessment, including population dynamics theory, length based stock assessment models, rebuilding analyses, estimation of biological reference points and harvest strategy modeling for invertebrates, as well as invertebrate biology. The CIE consultants travelled to Seattle, Washington to meet with the four member Interagency Work Group charged with developing the new overfishing definitions. One member of the Panel was present at the May meeting of the NPFMC Crab Plan Team in Seattle.

Description of Review Activities

AFSC provided access to a number of relevant papers that were listed on their web site www.afsc.noaa.gov/refm/stocks/CrabWs.htm and provided some additional documents by email. The key papers that focused on area of review were:

- Statement of work for working group.
- Description of proposed overfishing definition tier system.
- Stock assessments for Red King Crab and Snow Crab.
- Working group position papers.
- Workshop report recommendations.
- Projection model results.

A copy of the code for the snow crab stock assessment, and the AD Model Builder and FORTRAN code used for reference point estimation was offered to the review team but this was not required.

This CIE review team was asked to focus on:

a. A statement of the strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.
b. Recommendations for improvements to proposed overfishing definitions or alternative definitions,
c. A review of the model configurations, formulations and methods used to account for uncertainty.
d. A review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in simulation models.
e. Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

The panel met with scientists from the Alaska Fisheries Science Center and the Alaska Department of Fish and Game from April 24 to April 28, 2006, in Seattle, Washington. The meeting was chaired by Dr Anne Hollowed and Dr Jim Ianelli. The crab team presented the key aspects of their research on the first three days according to the agenda in Appendix 2. Throughout the presentations the CIE panel asked detailed questions on issues of the stock assessment and related research that was presented. All members of the crab team answered questions and expanded on some aspects of the stock assessment. On the fourth day the CIE panel met to highlight the key issues in the stock assessment modeling and overfishing definitions that would require some comment. They sought clarification from some members of the crab team on a number of issues before preparing to write their individual independent reports.

Summary of Findings

The findings of the review have been presented based according to the terms of reference set of the panel:

1. *A statement of the strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.*

Federal legislation requires an overfishing definition (OFD) that specifies whether the stock is overfished and whether there is overfishing occurring. The proposed OFD is a tier system that represents a significant improvement on the current system. The proposed system is based on the current NPFMC groundfish system which has been reviewed and hence provides a good basis for developing OFD. The groundfish system has incorporated a buffer between the overfishing limit (FOFL) and the target F level as required on the National Standard guidelines 1 (NSG1). In the current crab tier system there is no buffer between the target F and FOFL.

The proposed framework is comprehensive having five tiers which take into account the level of knowledge and uncertainty about the stocks being managed, i.e. whether reliable estimates are available for biomass and reference points and whether a stock assessment model has been implemented. However the uncertainty within a tier has not been thoroughly taken into account and should be considered when considering the overfishing and overfished definitions and the strategies for rebuilding. For Tiers 1 to 4 there are three levels of stock status with a corresponding target fishing mortality rate corresponding to the overfishing limit (FOFL).

The annual assessment of the stock provides for an annual revised estimate of the OFD levels with a revision of the model approach, the parameters of the model and the new year's data. This provides the 'best' indication of the status of stock. However this could also be viewed as a weakness of the proposed OFD approach in that the OFD can change with each year's stock assessment. There does not appear to be an assessment that compares the latest year's stock level and exploitation with the OFD level set the previous year for overfished and overfishing.

A two-stage approach should be considered for each year's stock assessment: (1) a comparison of the latest year's stock level and exploitation with the OFD level set the previous year definition for overfished and overfishing; and (2) undertake a revised stock assessment which may include a new model approach, revised biological parameters, new time series of data as well as the addition of the usual new year's data (such as survey, catch and effort). The changes to the previous years' assessment should be well documented and subject to review.

Modelling of the proposed overfishing tier system by the two modeling groups is viewed as a strength in the process of determining the OFD in that it provides a comparison of alternative approaches, different set of assumptions about the features in the model such as the measure of stock (B) which is the basis of the overfished assessment and the type of the stock-recruitment relationship (SRR). However to gain maximum benefit from the two modeling approaches it is important to undertake critical analysis of the results and provide a revision and improvement to the models. Some revision of the models has occurred but no consensus on the optimum model has been reached.

The projection model to compare rebuilding strategies should have the same starting biomass for each simulation. This was undertaken by Turnock and Rugolo (2006) but Siddeek and Zheng (2006) use a different starting value (beta x Bmsy) for some of the different models that evaluate the parameters. This means that the simulations are not comparable. Siddeek and Zheng (2006) have undertaken simulations to compare alpha and beta however because of the different starting values in biomass for different levels of beta, only alpha levels can be compared for different levels of beta. A range of starting values, eg .1-.7 Bmsy, should be used to test alpha and beta parameters. The different levels of alpha (0 to 0.1) tested show little difference in rebuilding time and long-term mean yield so any value in this range appears satisfactory. This is one of the weakness in the approach in the choice of alpha and beta are somewhat arbitrary and default levels of 0 and 0.2 can be used in the absence of evidence to indicate that there are more appropriate measures.

A weakness of the analysis is that there should be an assessment of the short-term impact of rebuilding on catch. There is no assessment of short-term impact on yield of the rebuilding strategies. This is usually one of the key elements of rebuilding that is required by managers and industry.

The projection model tests the harvest rule from the proposed Tier system as well as the current OFL and the current ADFG harvest strategies. The simulation confirms that the current OFL is not sustainable (Turnock and Rugolo 2006). Turnock and Rugolo (2006) provide a good comparison of a large number of rebuilding strategies including the F=0 and Fmsy strategies to help select the set of appropriate strategies. Siddeek and Zheng (2006) only focus on the OFL as the harvest strategy to test the rebuilding strategy which unnecessarily constrains the harvest strategy that may be required.

As you move down the Tiers 2 to 4, the models are more sensitive to scientist decisions as less information is available and hence require additional simulations to assess the relative merits of the model.

Tier 5 average catch may not be a conservative OFD depending on exploitation and recruitment patterns. Tier 5 should consider effort data in setting a target catch level. For example, has there been an increase and decrease in effort for the periods under consideration? If there is considerable annual variation in recruitment then this increases the chance of overfishing if there as a series of below-average recruitment. Simulation analyses associated with this Tier should be conducted to assist in determining a sustainable control rule. An initial OFL at a level below the average catch should be considered until there is evidence that the stock can support a higher catch.

A 3-year moving average of the levels in the overfished and overfishing definitions should be considered to assess the trends in the abundance and exploitation indices and reduce the possible biases in the annual indices. Therefore an average over 3 years will avoid the short-term impact of factors such catchability variability and assist in focusing the control rules on the significant trends in the fisheries.

2. *Recommendations for improvements to proposed overfishing definitions or alternative definitions,*

Some recommendations for improvements to the OFDs are described above. This section contains some additional recommendation to assess the OFDs:

- An assessment should be made of the short-term impact of rebuilding on catch. The trade-off relationship between rebuilding time and loss of short-term yield should be examined to determine an appropriate rebuilding time that minimises the short-term impact on the industry.
- There is a need to consider variability in the parameters, observation error, and hence the uncertainty associated with the current status relative to the decision rules within each of the tiers and the uncertainty associated with rebuilding

strategies so that managers can be aware of the variability associated with these assessments.

- A range of starting values, eg .1-.7 Bmsy, should be used in the rebuilding simulations to test alpha and beta to assess if there are more appropriate levels of alpha and beta than the arbitrary levels of 0 and 0.2.
- Additional simulations are required to assess the relative merits of the OFD models as you move down the tiers 2 to 4. These models are more sensitive to scientist decisions as less information is available. Tier 4 requires additional simulations to assess an additional parameter (gamma).
- Simulation analyses should be conducted with Tier 5 to assist in determining a sustainable control rule. An initial OFL at a level below the average catch should be considered until there is evidence that the stock can support a higher catch.

3. *A review of the model configurations, formulations and methods used to account for uncertainty.*

4. *A review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in simulation models.*

This section deals with Terms of Reference 3 and 4.

A measure of the egg production is a critical component of the population dynamics. This measure is particularly difficult for the Alaskan crab fishery which is a male only fishery resulting in a large numbers of mature females that are unmated, females with clutches that are not filled, females with unfertilized eggs, barren and senescent females. These are all indicators of a much lower abundance of mature males compared to mature females which results in the mature males being the limiting factor in the determining the egg production. There appears to be considerable annual variation in the fraction barren females and clutch fullness and it is important to understand the factors affecting this annual variation such as the effects of fishing and the environment. There is evidence that relates the level of exploitation (on the males) to the level of barren females, clutch fullness and females with unfertilized eggs.

Despite the harvest strategy with size limits set so that the males can mate at least once before being retained, the number of males still appear to be a bottleneck in the reproduction process. Hence the annual variation in the mature male abundance (taking into account sperm relationship with size) in the appropriate location may be the key determinant to egg production and should be considered as a possible indicator of egg production.

The current indicators being used for mature biomass in the OFD and the stock recruitment relationships do not appear good indicators of egg production and should be reviewed. The indicator used by Turnock and Rugolo (2006) takes into account the fact that mature males are limited in determining effective mature female biomass but then it adds the effective male mature biomass which does not appear appropriate.

The cause of the reduction in the king crab stocks since the 1980's is critical in determining what are the target Bmsy levels. If the reduction is due to a regime shift then basing the Bmsy on the lower levels of mature biomass since the 1980's is appropriate. Dew and McConnaughey (2005) provide evidence of the negative effects of the increase in trawling in 1980, particularly in the most productive spawning grounds off Unimak and Amak Islands, on the breeding stock. This impact would be exacerbated if the area is correctly identified as a valuable 'source' area and contains high abundance of multiparous crabs. The highly aggregated behaviour of the king crabs further increases their susceptibility to overfishing. Even if the reduced biomass is due to the effects of trawling, it may not be possible to restrict trawling from the more productive spawning areas and re-introduce the appropriate sanctuary zones. In this case basing the Bmsy on the lower levels of mature biomass since the 1980's is still appropriate as the breeding stock will not return to the levels of the 1970's under the current levels of trawling. However if the impact on the trawling on the spawning biomass can be reversed then basing the Bmsy on the level of mature biomass of the 1980's may significantly underestimate the true potential of the stock. An adaptive management approach should be considered to assess the effects of fishing on these productive grounds by closing an appropriately-sized area to trawling to determine the impact on the mature stock in that area.

The two competing hypotheses on decline of the king crab stocks since the 1980's, i.e. regime shift and the effects of increased targeted and trawling, may both be contributing to the decline in recruitment. Many stocks quite often collapse when there is the combined effect of poor environmental conditions at a time when the breeding stock is reduced to changes in fishing practices.

The relationship between male molting and subsequent mating of snow crab has been a source of different interpretations between the research teams. While after the males molt, they '**can potentially mate** with primiparous females the following winter and with multiparous females in the spring of the following year', however the newshell males are **outcompeted as mates** (Workshop report, 2006). If these males are used as contributors to the egg production (Zheng 2006) then they should be discounted to reflect the biological qualifications associated with the mating contribution by these males.

As the relative size of mature males and females is import in the mating process, it is important to monitor the changes in mean size and length frequency for mature males and females that occur. The ratio of mature male to mature female mean size could also be used to measure the relative changes in mean size.

The choice of the stock-recruitment relationship (SRR) is important in the stock assessment of the Alaskan crab fisheries and both modeling groups have given this issue a significant level of attention. The Maximin Clark (1991) method provides a basis to assess different steepness levels of the SRRs when there is no empirical data available. However in many cases there are some data available to at least make a choice about whether the SRR is likely to be a Ricker or Beverton-Holt curve. This would at least restrict the choices available and result in a more appropriate choice. This empirical data can also be used in the development of informed priors, eg relative probabilities Beverton-Holt and Ricker curve, when in the stock assessment models. Siddeek provided a valuable assessment on the relationship between Tau and steepness in the SRR of the Ricker and Beverton-Holt curves.

The SRR is affected by the years chosen to assess the fit. There is a significant change to the recruitment pattern before and after 1976. Irrespective of whether this change is due to a regime shift or the effects of trawling, there will be a change in the shape of the SRR and this should be taken into account. The change in shape of the SRR may take the form of a stock-recruitment-environment relationship (SRR-E) which takes into account the regime shift or the effect of wind on the recruitment of Tanner crabs (Rosenkranz et al. 1998). Even if the reduction in recruitment is due to the effects of fishing, then a dummy variable can be used in the SRR to differentiate the years before and after 1976.

The Turnock and Rugolo (2006) population models have a large number (276) of parameters estimated and it appears these could be significantly reduced. For example, there appears to be little biological basis for having separate male and female recruitment indices (even if they 'were constrained to be similar'). The annual recruitment of males and females should be similar and set at appropriate sex ratio if the recruitment sex ratio is not 1:1.

The biological basis for having different selectivities for new and old shell is not clear (Fig 20 and 21 in Turnock and Rugolo 2006). Annual parameters are estimated for selectivities and again it is not clear why selectivity should change every year. In fact Figure 21 indicates that selectivity for new shell appears constant over the years and hence the number of parameters could be reduced. There appears to be a dramatic difference in the shape of the survey selectivity before and after 1982 (Fig. 22 in Turnock and Rugolo 2006) with an increase in selectivity for the larger sizes and decrease in selectivity for smaller crabs. However the reason for this change in selectivity is not explained.

The use of different natural mortality levels for 3 different periods for males and 4 different periods for females (Zheng 2006) does not appear to be biologically sensible. While it is possible for mortality to vary over the years it does not appear to be reasonable for the differences to be at different times for the sexes. The application of different levels of mortality appears to be based on the statistical fit of the model which could be explained by a number of reasons of which variation in natural mortality is only one possibility.

Estimation of survey catchability for snow crabs using underbag have been undertaken. However this may not provide a complete assessment of the catchability. The use of a depletion experiment should be considered to estimate survey catchability for different sizes, shell condition and sexes.

Environmental factors can have a significant impact on the efficiency of the gear and it would be useful to have an assessment of this issue. The key environmental indices during the surveys should be summarized so that the potential biases in the indices are identified and whether that bias is likely to be positive or negative. If the relationship between the environmental factors and gear efficiency can be determined then this relationship can be used to standardize the catch rates so that they better reflect the abundance of the year-classes.

5. *Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.*

    a. A measure of the egg production is a critical component of the population dynamics. This measure is particularly difficult for the Alaskan crab fishery which is a male only fishery resulting in a large numbers of mature females that are unmated, females with clutches that were not filled, females with unfertilized eggs, barren and senescent females. These are all indicators of a relatively lower abundance of mature males compared to mature females which results in the mature males being the limiting factor in the determining the egg production. Hence the annual variation in the mature male abundance may be the key determinant to egg production and should be considered as a possible indicator of egg production. The current indicators being used for mature biomass in the OFD and the stock recruitment relationships do not appear good indicators of egg production and should be reviewed. An adaptive management approach should be considered to assess the effects of trawling on the previously productive breeding grounds off Unimak and Amak Islands by closing an appropriately-sized area to trawling to determine the impact on the stock in that area.
    b. Depletion experiments should be considered to estimate survey catchability for different sizes, shell condition and sexes.
    c. A depletion analysis of some blocks that are heavily fished during a season such that there is a significant decline in catch rate due to the effects of fishing could provide some valuable insights into some fishery dynamics. A comparison of the daily retained male CPUE in a block (or groups of blocks) and the cumulative legal catch removed from that block over the period that the fishery operates enables an estimate of the residual legal biomass at the end of fishing, the catchability of the crabs and the exploitation rate.
    d. A depletion analysis may also be applied to assess the impact of fishing on discards if there is sufficient observer data on the daily catch rate of discards in a heavily fished block(s) and an estimate of discard numbers can be made from those block(s). A significant decline in the discard rate during the course of fishing would indicate a significant level of discard mortality.
    e. The change in the management of the fishery to an individual transferable quota (ITQ) is likely to result in high grading and hence increase the rate of discarding and associated discard mortalities. Consideration should also be given to retaining some of the discards by providing a separate quota for discards. If there is a high mortality (50-100%) associated with discards it

may be worth retaining some of them (if there a market for them) and reducing the ITQ for the first-grade crabs.

f. While considerable research on escape gaps and subsequent changes have been undertaken on escape gaps, it appears that there is still considerable retention of undersize crabs, most (50-100%) of which may die as a result of being captured. This makes it imperative to undertake further research (if necessary) to choose the number and size of the escape gaps that maximizes the escape of undersize male and female crabs even if it means that some of the smaller legal-size males are allowed to escape. Additional research on the handling practices (dropping crabs on a hard surface from a height of greater than 4 ft) onboard should also be undertaken to assess if there are ways to improve handling practices to increase survival of discards.

g. An evaluation should be undertaken on the merits of retaining some female king crabs that are marketable as part of the catch. There appears to be a surplus number of mature females relative to the number of mature males in the fishery resulting in unmated and senescent females. These females could contribute to significant loss of productivity due to density dependent mortality and growth, particularly if habitat is limiting. The discarding of female crabs results in a high discard mortality in which case there appears to be a significant wastage of product. The retention of an approved quantity of females would provide a basis for increasing the overall yield or can be used to offset a reduction a male catch and hence result in an optimum sex ratio for mating. A modeling of harvest strategy should be examined that includes the retention of an appropriate quantity of females that results in an optimum ratio of mature males to mature females and hence a reduction in unmated mature females.

h. The modeling of the shell condition is a critical part of the population dynamics of the crab fishery as it affects the catch that is targeted and retained, molting, growth, maturity and the mating dynamics. There appears to be uncertainty about the relationship that has been assumed between shell condition and time since last moulting and this relationship needs to be examined further.

i. An economic assessment of the fishery should be undertaken in conjunction with the stock assessment modelling to assess ways to improve the economic performance of the fishery. The maximum economic yield (MEY) which is less than MSY should be considered as a performance indicator for the fishery as it would be a more conservative indicator.

j. An assessment should be made of the short-term impact of rebuilding on catch. The time trend in rebuilding of biomass has been presented by Turnock and Rugolo (2006). Trade-off relationship between rebuilding time and loss of short-term yield should be examined to determine an appropriate rebuilding time that minimises the short-term impact on the industry. This information is vital for economic analysis of any rebuilding strategy.

References

Dew, C.B. and McConnaughey, R (2005). Did trawling on the brood stock contribute to the collapse of the Alaska's king crab? Ecological applications 15: 919-941.

Rosenkranz, G, Tyler, A. Kruse, G. (2001). Effects of water temperature and wind on year-class success of Tanner crabs in Bristol Bay, Alaska. Fisheries Oceanography 10: 1-12.

Siddeeek, S., Zheng, J. (2006). Reference point estimation analysis for the Bering Sea and Aleutian Islands (King and Tanner) crab revised fisheries management plan. Alaska Department of Fish and Game (draft).

Turnock, B., Rugolo, L. (2006). Stock assessment of eastern Bering Sea snow crab. National Marine Fisheries Service (draft).

Workshop Report (2006). Crab overfishing definitions inter-agency workshop. February 28 – March 1, 2006.

Zheng, J. (2006). Bristol Bay red kink crab stock assessment in 2004. Alaska Department of Fish and Game (draft).

Appendix 1

Consulting Agreement between the University of Miami and Dr. Nick Caputi

STATEMENT OF WORK

April 27, 2006

Background

The Alaska Fisheries Science Center (AFSC) requests review of proposed overfishing definitions and simulation models used to evaluate biological reference points for Bering Sea and Aleutian Islands King and Tanner crab stocks. The North Pacific Fishery Management Council (NPFMC) has determined that the existing overfishing definitions for Bering Sea and Aleutian Islands King and Tanner crab stocks need revision. The AFSC is seeking review of the population dynamics models developed for revising the overfishing definitions.

There are currently 22 Bering Sea and Aleutian Islands crab stocks under the Federal Bering Sea Aleutian Island Crab Fishery Management Plan (FMP) of which 7 are considered major stocks. Four of the seven major crab stocks have been declared overfished and rebuilding plans developed within the last 7 years. Of the remaining three stocks, only one has been relatively stable at a low level, another has maintained stable catch for several years, however, even for this stock it appears recruitment may be declining. While the remaining stock has increased, survey abundance estimates have low precision and the fishery is closed due to bycatch concerns. There is no consensus on the principal cause of declines in Bering Sea crab stocks.

Review Requirements

A panel of three consultants is requested for this review. In aggregate, the panel will need to be thoroughly familiar with various subject areas involved in the review: crab biology; analytical stock assessment, including population dynamics theory, length-based stock assessment models, rebuilding analyses, estimation of biological reference points and harvest strategy modeling for invertebrates; and AD Model Builder. The CIE consultants will travel to Seattle, Washington to meet with the Interagency Work Group charged with developing the new overfishing definitions. We request that one member of the Panel should be present at the May meeting of the NPFMC Crab Plan Team in Anchorage, Alaska. We also request that one member of the Panel be present at the June meeting of the NPFMC Scientific and Statistical Committee meeting in Kodiak, Alaska. It would be preferable that the same individual attends both of these meetings, but this is not a requirement.

The report generated by each consultant should include:

f. A statement of the strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.

g. Recommendations for improvements to proposed overfishing definitions or alternative definitions,
h. A review of the model configurations, formulations and methods used to account for uncertainty.
i. A review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in simulation models.
j. Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

AFSC will provide copies of the NPFMC Work Group statement of work, proposed overfishing definitions, preliminary results of simulations, discussion of input parameters, a copy of the code for the snow crab stock assessment, and the AD Model Builder and Fortran code used for reference point estimation. The panel will meet with scientists from the Alaska Fisheries Science Center and the Alaska Department of Fish and Game from April 24 to April 28, 2006, in Seattle, Washington (see attached agenda).

It is estimated that the duties of each reviewer will occupy a maximum of 14 days each: several days for preparation, five days for the workshop, several days for writing their reports, and two days for travel. In addition, a maximum of nine reviewer days will be allowed for attending the two council meetings, including preparation time, travel, and one day to attend each meeting. The total level of effort is 51 days of reviewer time.

Products

- One member of the panel will attend the May meeting of the Crab Plan Team on May 17, 2006 in Anchorage, Alaska, to discuss the panel's findings regarding the strengths and weaknesses of proposed definitions and modeling approaches.
- One member of the Panel will attend the June meeting of the NPFMC Scientific and Statistical Committee meeting on June 5, 2006 in Kodiak, Alaska, to discuss the panel's findings regarding the strengths and weaknesses of proposed definitions and modeling approaches.
- No later than May 12, 2006, each panelist shall submit a written report of findings, analysis, and conclusions. See Annex 1 for details on the report outline. The reports should be sent via e-mail to Dr. David Die at ddie@rsmas.miami.edu, and to Mr. Manoj Shivlani at mshivlani@rsmas.miami.edu.

Appendix 2 Meeting Agenda

Center of Independent Experts
Alaska Crab Overfishing Definitions

April 24 - 29, 2006

Alaska Fisheries Science Center, Seattle, WA

Purpose: To solicit expert advice on proposed overfishing definitions for Bering Sea and Aleutian Islands crab stocks. We are requesting a review of issues critical to formulating new overfishing definitions, biological reference points, input parameters, modeling approaches and methods to deal with uncertainty.

DAY 1 (Center Director's Conference Room)

8:00 Coffee and informal discussions

8:30 Introductions - Charge for the CIE –Hollowed

8:50 History of crab management - current overfishing definitions and need for revision - Stram or Designee

9:10 Overview of proposed revisions - Working group

- Working group Statement of Work (20 min) - Rugolo
- Tier System review (20 min) - Zheng
- Brief Description of Snow Crab Assessment (40 min) -Turnock

10:30 Break

10:30 – 12:00 Overview continued – working group

- Brief Description of Red King Crab Assessment (40 min) -Zheng
- Projection Model structure (Siddeek and / or Turnock)

12:00 – 1:00 Break for lunch

1:00-1:30 Overview continued – working group

- Approaches to estimate proxy values for biological reference points – Turnock
- Approaches to estimate proxy values for biological reference points - Siddeek

1:30 – 2:00 Review Workshop Report recommendations on crab biology – Stram or designee

2:00 – 2:30 Review of Workshop Report recommendations on crab modeling - Ianelli

2:30 Break

2:45-3:45 Review of information available for managed crab stocks - Rugolo

3:45 – 5:00 Performance of Tier System Preliminary results

- Red King Crab – Siddeek
- Red King Crab – Turnock

DAY 2 (CD Conference Room) 8:30 Coffee and informal discussions

8:30 – 10:00 Performance of Tier System Preliminary results continued

- Snow Crab – Turnock
- Snow Crab – Siddeek
- Blue King Crab/Golden Crab - Siddeek

10:00 Break

10:30 – 12:00 Questions and Answers for panel.

12:00 Lunch

1:00 – 5:00 Open question and answer session – or independent work sessions with CIE reviewers.

DAY 3 (CD Conference Room)

8:30 Coffee and informal discussions

9:00 Open question and answer session – or independent work sessions with CIE reviewers.

DAY 4 (CD Conference Room)

8:30 Panel discussions and writing team – NMFS and ADF&G biologists return to offices but remain on call to answer questions

DAY 5 (CD Conference Room)

8:30 Panel discussions and writing team – NMFS and ADF&G biologists return to offices but remain on call to answer questions

APPENDIX 3: Bibliography of materials provided during the review meeting

The key documents referred to during the review are listed below:

Dew, C.B. & McConnaughey, R.A., 2005. Did trawling on the brood stock contribute to the collapse of Alaska's king crab? *Ecological Applications*, **15**, 919-941.

Maunder, M.N., 2003. *Review of the stock assessment and harvest strategy for eastern Bering Sea snow crab.* CIE, University of Miami.

NPMFC, 2006. *Workshop Report: Crab Overfishing Definitions Inter-agency Workshop. February 28-March 1, 2006, Alaska Fisheries Science Center, Seattle, WA.* NPMFC, Anchorage.

Restrepo, V.R., Thompson, G.G., Mace, P.M., Gabriel, W.L., Low, L.L., MacCall, A.D., Methot, R.D., Powers, J.E., Taylor, B.L., Wade, P.R. & Witzig, J.F., 1998. *Technical guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act.* NOAA Technical Memorandum NMFS-F/SPO-##.

Rugolo, L., 2004. *North Pacific Fisheries Management Council Bering Sea/Aleutian Islands King and Tanner Crab Working Group: Draft Statement of Work*. NMFS/ADF&G, Kodiak/Seattle/Juneau.

Rugolo, L. 2006. *Statement of Work: NPFMC BSAI King and Tanner Crab Working Group*. www.afsc.noaa.gov/refm/docs/2006/crab/Statement%20of%20Work.ppt

Siddeek, M.S.M. & Zheng, J., 2006. *Reference point estimation analysis for the Bering Sea and Aleutian Islands (king and Tanner) crab revised fisheries management plan.* ADF&G, Juneau.

Turnock, B.J. & Rugolo, L.J., 2005. *Stock assessment of eastern Bering Sea snow crab.* NMFS, Seattle/Kodiak.

Turnock, B.J. & Rugolo, L.J., 2006a. *Analysis of proposed overfishing tier system for BSAI king and Tanner crab stocks.* NMFS, Seattle/Kodiak.

Turnock, B.J. & Rugolo, L.J., 2006b. *Unresolved issues concerning proposed overfishing definitions for Bering Sea and Aleutian Islands king and Tanner crab stocks: National Marine Fisheries Service.* NMFS, Seattle/Kodiak.

Zheng, J., 2004. *Bristol Bay red king crab stock assessment in 2004.* ADF&G, Juneau.

Zheng, J., 2006. *Issues dividing the Crab Work Group.* ADF&G, Juneau.

Zheng, J. & Kruse, G.H., 2006. Recruitment variation of eastern Bering Sea crabs: Climate-forcing or top-down effects? *Progress in Oceanography*, **68**, 184-204.

Further documentation available to the reviewers, including presentations given to the crab overfishing workshop is given at:

http://www.afsc.noaa.gov/refm/stocks/CrabWs.htm

REPORT ON

BERING SEA & ALEUTIAN ISLANDS

CRAB STOCK OVERFISHING DEFINITIONS

24-27 APRIL, 2006

SEATTLE, WASHINGTON

Prepared by

Patrick Cordue
Fisheries Consultant
New Zealand

for

University of Miami
Independent System for Peer Review

6 June 2006

EXECUTIVE SUMMARY

A CIE Review Panel considered a proposed overfishing definition for Bering Sea and Aleutian Islands crab stocks from April 24-27, 2006 at Alaska Fisheries Science Center, Seattle, WA. The existing definition had been found to be in need of revision and an interagency work group had been charged with developing a new definition. They had encountered difficulties in doing this and a two-day workshop had been held to discuss and resolve issues. The CIE review took place about 8 weeks after the workshop. In the interim, the work group had continued working on the overfishing definition framework. In particular, they had attempted to find suitable default parameter values and proxies needed to complete the overfishing definition.

The proposed overfishing definition is an improvement on the existing definition in that it provides some constraint on fishing mortality. The existing definition is flawed in concept, does not constrain fishing mortality, and needs to be replaced.

The proposed definition is:

- an improvement on the existing definition
- comprehensive (as a framework)
- borrowed from groundfish (so is already reviewed to some extent).

Weaknesses of the proposed definition:

- complicated
- it is still a work in progress
    - default values for parameters are not yet determined
    - sensible definition of biomass in the stock recruit relationship is not determined/specified
    - criteria for determining optimal default parameters are not determined/specified
- extensive simulations are needed to determine suitable default parameters
- potentially, it may unnecessarily constrain harvest strategies.

I make several recommendations. The most important of these concern two central issues: the definition of biomass in the stock recruit relationship, and the criteria for choosing between overfishing-definition MSY control rules.

The issue of the definition of “biomass” in the stock recruit relationship is peculiar to crabs because fishing mortality is only directed at males. In groundfish stocks it is not an issue because female spawning biomass is a good proxy for total fertilized egg production. For crabs it is a crucial issue for the proposed overfishing definition because the biomass proxy for total fertilized egg production is a primary determinant of F_{MSY} and F_{MSY} proxies. To date, the analysis of this issue has been inadequate. Immediate efforts need to go into the derivation of appropriate functional forms. In the short term, if a default definition is needed, mature male biomass should be seriously considered.

Also, there is the issue of what constitutes a "good" overfishing definition, in general, and for Bering Sea Aleutian Islands crabs in particular. The answer to this question needs to be clearly stated. It is then relatively straightforward to define the analysis and simulations needed to test alternative overfishing definitions (and to determine default parameter values for the MSY control rules in the proposed tier system). The function of an MSY control rule in an overfishing definition must be acknowledged. The preliminary simulations aimed at determining default parameter values tested MSY control rules as rebuilding plans and harvest strategies. They are neither. MSY control rules must be evaluated in conjunction with harvest strategies (either existing harvest strategies, or a default harvest strategy).

The parameterization of the proposed MSY control rules implies a reduction in F at B_{MSY}. It does allow flat control rules (alpha = –infinity) but it precludes the suggested default overfishing definitions of Restrepo et al. 1998 (where the reduction in F occurs below B_{MSY}). I suggest that an extra parameter is added to the framework to allow MSY control rules of the form proposed by Restrepo et al. (1998). In the absence of this parameter, the proposed framework may unnecessarily restrict harvest strategies.

BACKGROUND

A three person CIE Review Panel considered a proposed overfishing definition (OFD) for Bering Sea and Aleutian Island (BSAI) crab stocks from April 24-27, 2006 at Alaska Fisheries Science Center, Seattle, WA. The North Pacific Fishery Management Council had determined that the existing OFD needed revision. A four member interagency work group had been charged with developing the new OFD. They had already participated in, and taken direction from an interagency workshop on crab OFDs which had met February 28-March 1, 2006. Simulation studies, aimed at determining default parameter values and proxies needed in the proposed OFD framework, were undertaken between the OFD workshop and the CIE review meeting.

This report presents my personal view with regard to the proposed OFD and the methods and techniques needed to determine appropriate default parameter values and proxies. I also comment on the stock assessment models and estimation methods in general. Finally, I suggest some research priorities. This report should be read in conjunction with those of my fellow reviewers Dr Mike Bell and Dr Nick Caputi. Although there was no attempt to reach a consensus on any of the issues it was apparent that the Review Panel shared many common views with regard to the proposed OFD and associated research.

REVIEW ACTIVITIES

Meeting Preparation

Prior to the meeting I read the main documents and consulted the background material made available on a website (Appendix 1). I also consulted material on the Web and conversed with colleagues with regard to crab biology.

Meeting Attendance

A brief narrative of the meeting is given below.

24 April

The meeting was convened at 8.30 am and began with a round of introductions. The meeting Chair, Dr Anne Hollowed, gave an introductory presentation on the purpose of the review and the "charge for the CIE". Dianna Stram reviewed the history of crab management and the existing OFD and the reasons for revision. Simply put, the existing OFD had been rushed through; it was conceptually flawed and provided no constraint on fishing mortality.

The four member Working Group then covered material relating to their statement of work, that of the two-day workshop, the proposed OFD structure (tier system and parameters) and two example stock assessments (snow crab and red king crab).

The Review Panel asked many questions during the presentations. We were aware that slow progress was being made in terms of the original agenda but thought that it was best to fully explore the issues during the presentations. We had already advised the Chair that we would not need to use the whole week. The scheduled "writing team" days were not needed as Panel members agreed that we could best do this after returning to our home locations.

25 & 26 April

The meeting resumed at 8.30 am with Dr Jim Ianelli in the Chair. We began with a presentation on the projection model structure (Dr Siddeek). This was followed by a presentation on approaches for estimating F_{MSY} and B_{MSY} proxies (Dr Turnock). The report on the interagency workshop (Anon. 2006) was reviewed briefly since we had already discussed most of the issues considered in it.

During the rest of the day and during the next day, preliminary simulation results were presented by the Working Group members. Attempts had been made to evaluate different alpha, beta, and gamma parameter values. Also, some proxies for F_{MSY} had been tested. However, in all cases the results were preliminary and no firm recommendations could properly be made with regard to proxies or default parameter values on the basis of the simulations.

27 April

The Review Panel convened at 9.30 am to identify, discuss, and clarify all relevant issues relating to the proposed OFD and to supporting research. We covered points a.-e. as per our Statement of Work (Appendix 3). Late in the day we had a question and answer session with Dr Hollowed, Dr Turnock, and Dr Rugolo.

Post Meeting Activities

Prior to and during my return journey to New Zealand I considered the two main problems that the Working Group were grappling with.

First, they had not fully defined the criteria for choosing between alternative tier-structure parameter values (in terms of being the best defaults). This, I believe, stemmed from the fact that the problem had not been fully specified. In order to determine the best defaults, one must define what it is for one MSY control rule to be better than another when they are used as part of an OFD.

Second, there had been inadequate analysis used to define "biomass" (B) in the stock recruit relationship (SRR). The Working Group had found that their results were very sensitive to the definition of B. They did not have an adequate definition and had no means of choosing between the alternatives they had proposed. I spent considerable time exploring alternatives for deriving appropriate functional forms – the aim being to illustrate how total fertilized egg production could be expressed as a function of population parameters (which could conceivably be measured or estimated).

The lead reviewer, Dr Bell, was to present our findings at two meetings which were scheduled earlier than the original deadline for production of our reports. On my return to New Zealand I produced an interim report for Dr Bell, in advance of his first meeting, which, while short on detail, differed little in the conclusions and recommendations of this report. I also undertook to produce my final report well in advance of Dr Bell's second meeting (but some days after the new deadline specified in the revised SOW – see Appendix 3).

SUMMARY OF FINDINGS

The existing OFD is conceptually flawed and as a consequence places no constraints on fishing mortality. It clearly needs to be replaced, but care must be taken to ensure that its replacement does not overly constrain potential harvest strategies.

To my mind, there are two central issues to consider with regard to the proposed OFD.

First, there is the issue of what constitutes a "good OFD", in general, and for BSAI crabs in particular. If the answer to this question is clearly stated it is relatively straightforward to define the analysis and simulations needed to test alternative OFDs (and to determine default parameter values for the MSY control rules in the proposed tier structure). Related to this issue is the question of whether an OFD MSY control rule can be appropriately tested in isolation from a harvest strategy (HS). In reality, the MSY control rule imposes constraints on the HS which is used and so management strategy evaluation must test MSY-control-rule:harvest-strategy pairs.

The second central issue is the definition of B in the SRR. This issue is peculiar to crabs because fishing mortality is only directed at males. In groundfish stocks it is not an issue because female spawning biomass is a good proxy for total fertilized egg production (TFEP). This is a crucial issue for the proposed OFD because the biomass proxy for TFEP is a primary determinant of F_{MSY} and F_{MSY} proxies.

What constitutes a good OFD?

We should first consider the question, exactly what is an OFD? We should distinguish between an OFD for a particular stock and an "OFD framework" which specifies a family of OFDs. It is the latter which the review is concerned with and the "family" consists of OFDs for BSAI crab stocks. Central to an OFD is the concept of an MSY control rule, which defines F_{OFL} as a function of biomass and from which derives the overfished threshold (MSST). A "good OFD" (framework) can sensibly be defined as one which specifies "good MSY control rules".

The proposed OFD has a five level tier structure to accommodate stock assessments with different levels of reliability (Anon. 2006, page 8). The fifth tier is for stocks which are not formally assessed. In the first four tiers a linear parameterized MSY control rule is specified. F_{OFL} is constant above B_{MSY} and set equal to F_{MSY} or a proxy. Below B_{MSY} there is a linear reduction in F_{OFL} governed by two parameters alpha and beta. In tier 4, the F_{MSY} proxy is the product of the parameter gamma and M. The fishery is closed when estimated biomass (as a proportion of B_{MSY} or its proxy) is less than beta.

As it stands, the OFD appears incomplete until some default parameter values and proxy definitions are specified. In order to do this, criteria must be specified for determining when one MSY control rule is better than another. Given the criteria, alternative parameter values and proxies can be tested by doing model simulations over an appropriately broad range of population models (i.e., with different biological parameters and/or SRRs and/or model structures; the range being appropriate to the tier being tested).

The criteria for determining whether one MSY control rule is better than another were not discussed during the review meeting. From the preliminary simulations it appears that the implicit criteria relate to their performance as rebuilding strategies (since simulations were done from starting values less than BSST or at beta, with catches set at the OFL). The ranking of MSY control rules on the basis of their performance as rebuilding plans, or more generally as harvest strategies, is inappropriate given that is not their *function* (when specified in an OFD). The function of an OFD MSY control rule is to constrain (estimated) fishing mortality and to provide Status Determination (i.e., MFMT and MSST). It impacts on whatever harvest strategy is used for setting OY but it is not the harvest strategy (or the rebuilding plan).

I note that simulations using an MSY control rule as a harvest strategy are required to determine MSST (Restrepo et al. 1998). This is because the full definition of MSST is the maximum of two values: half B_{MSY} and "the minimum stock size at which rebuilding to the MSY level would be expected to occur within 10 years if the stock or stock complex were exploited at the maximum fishing mortality threshold". (During the review meeting no such simulations were discussed and it was (implicitly) assumed that MSST always equaled half B_{MSY}. In general, this should not be taken for granted.)

Restrepo et al. (1998) offer some advice on choosing an MSY control rule. Two factors are mentioned. First, the position of MSST may be of interest in that a council could "minimize the range of stock sizes within which special rebuilding plans would be required" if it opted "for an MSY control rule that afforded a good deal of 'built-in' rebuilding". The proposed OFD has such MSY control rules in that the linear decrease in F_{OFL} begins at B_{MSY} (which is even more conservative than the default MSY control rule suggested by Restrepo et al. 1998). Second, they suggest that the "tradeoff between magnitude of yield and constancy of yield" could be used. This involves testing the MSY control rules as harvest strategies. As already discussed this is inappropriate since that is not their function in an OFD setting.

In practice, an (OFD) MSY control rule is never used as the harvest strategy. Councils are required to "adopt a precautionary approach to the specification of OY" (Restrepo et al. 1998). Obviously, from a management strategy evaluation perspective, MSY control rules cannot be tested in isolation. They must be tested with an associated harvest strategy.

The choice of an MSY control rule is primarily a management decision. The tradeoff is between potential yield and risk. If an MSY control is too constraining on harvest strategies it may unnecessarily reduce long term yield. Conversely, if it is too liberal it may allow harvest strategies which are not precautionary. Given the current requirement for an (OFD) MSY control rule and a precautionary harvest strategy, constrained by the MSY control rule, it is necessary to test MSY-control-rule:harvest-strategy pairs.

Once this conclusion is reached it becomes a matter of detail on how to determine appropriate defaults to complete the specification of the proposed OFD framework (or to justify a choice of OFD for a particular stock assessment). Any existing harvest strategies are candidates to be tested. In their absence I suggest adopting some "default" harvest strategies in the simulations (e.g., those derived from the MSY control rule by applying 75% of the estimated OFL in each year).

When doing such simulations it is important to distinguish each of the individual components. There is the "operating model", which is the model of reality, within which everything is known exactly (e.g., B_0, F_y for each year y, $F_{MSY,}$ etc). There is also the "estimation frame" where quantities are estimated, such estimates being a function of the truth (from the operating model) and error (e.g., an estimate from a stock assessment). When evaluating an OFD MSY control rule there must also be a HS. The role of the MSY control rule is limited. It defines F_{OFL} for any given biomass *estimate* and it defines whether the stock is overfished or not (on the basis of the biomass *estimate*). On the other hand, the HS is used to set the TAC in each year that an assessment is conducted (within the simulation model). There is a requirement for a buffer between the OFL and the TAC. Hence, simulations using the MSY control rule as the harvest strategy (i.e., $F_y = F_{OFL}$ in every year y) are entirely inappropriate.

Definition of B in the SRR

The definition of B in the SRR is of crucial importance in obtaining a precautionary OFD. Since the directed fishing mortality is only on males, females suffer fishing mortality only as incidental bycatch (and subsequent handling/discard mortality). If the usual groundfish definition for B, of total female spawning/mature biomass, is used then F_{MSY} and proxies for F_{MSY} (such as $F_{50\%}$) are very large in an absolute sense. Crab biology is such that the role of males is crucial in the production of fertilized eggs and it is clear that the males must be brought into the definition of B.

In the long term, a suite of deterministic population dynamics models should be derived specifically for crab stocks, taking account of the important role played by males in the SRR. In the interim, it is probably best to derive an appropriate functional form for TFEP and simply assume that mean recruitment is a Beverton Holt or Ricker function of TFEP.

The review material contains several alternative proposed definitions for B. Total female mature biomass was, I assume, used for illustrative purposes only. Total male and female mature biomass was put forward as a candidate. This must be rejected because as male biomass approaches zero, TFEP approaches zero, but total mature biomass does not. There were at least two variations of female mature biomass scaled down by an "effective fertilization factor" (derived from an assumed "mating ratio"). The concept behind these definitions is that TEP is proportional to female biomass and that successful fertilization depends on the proportion of mature males in the mature population and the average number of females that each male can mate (the "mating ratio" which is assumed to be constant).

The concept of a "mating ratio" is sound in principle. In practice, it was found that F_{MSY} and F_{MSY} proxies were sensitive to the assumed mating ratio. So, even if one of the proposed formulations was accepted it still leaves the problem of determining an appropriate parameter range for the mating ratio.

During the review I questioned the validity of the assumption that TEP is proportional to mature female biomass. Dr Rugolo presented results from trawl survey and experimental data on the total number of eggs per female as a function of clutch fullness, shell condition, mating category, and carapace width. It is known that older females tend to have lower clutch fullness and that crabs with very old shell condition (4 & 5) tend to be barren. This is a problem for the proportionality assumption in that increasing biomass (with age) is inversely proportional to EP. Though, if the proportion of older females stays relatively constant it may not of itself be a major problem. However, it was also indicated that clutch fullness is strongly related to mating category, at least in snow crabs, with primiporous females typically having a0.75 clutch fullness and first time multiporous females typically having full clutches. Further, within clutch fullness category, the number of eggs appeared to be linearly related to the carapace width. While these data cast considerable doubt on the biomass proportionality assumption, their existence provides the very means by which to construct a sensible functional form for TEP and possibly to estimate a mating ratio.

I have undertaken some preliminary work on the derivation of a suitable equation for TFEP (Appendix 2). This work is illustrative and not definitive. An experienced mathematician should work with crab biologists to derive appropriate forms (at different levels of complexity) for TFEP. I also indicate how the trawl survey data (available since 1995) could be used to estimate unknown parameters, including a mating ratio, within the equation for TFEP (Appendix 2). In the absence of this sort of work (i.e., given time constraints), the best proxy for TFEP may be total mature *male* biomass (TMMB).

This suggestion was made by Dr Caputi at the review meeting and at the time, after discussion, was considered to be deficient in that it was inappropriate for stocks near their virgin level. It was considered that we needed a relationship which would be sensible over the full range of stock sizes. In a severely depleted stock, it is clear that sperm availability is the determining factor in fertilization success (since there are plenty of females). It is reasonable to argue that TMMB could be approximately proportional to TFEP when a stock (through removal of males) is depleted below some level. The effective mating ratio doesn't need to be known – the assumption is made that there are always enough females and that the mating ratio is constant. Of course, above some level of TMMB the proportionality assumption must fail. In Appendix 2 I have suggested an appropriate functional form to adjust for this effect. It adds an extra level of complexity to the SRR but will be more realistic than assuming full proportionality.

Other issues

Most other issues are minor in comparison to the two central issues already discussed. However, they are numerous and potentially time consuming in their detail. Below, I give some general comments on some of the issues.

Having two modeling groups is both a strength and a weakness. The exchange of ideas is valuable. The natural competition which arises can be stimulating and lead to improved methods and models. However, differences in modeling approaches can become entrenched; argument rather than discussion can be the outcome. While all members of the Working Group were cordial, helpful, and professional during the review meeting, there was clearly some tension between the two groups. In New Zealand, "contested stock assessment" is a common feature of our annual stock assessment cycle (Starr et al. 1998). We have recently agreed on some principles to help competing modeling groups work together:

- consider all components of the models and estimation procedures
- identify where differences exist between the two approaches
- where there is a "best" way to do something, agree to do it that way
- where there are two (or more) reasonable alternatives, implement both (all) options
- ideally, each modeling group should be able to reproduce the results of the other group (but, if totally different estimation procedures are being used, this is probably not an option).

Difference results do not present a problem if the reason for the differences is understood. In New Zealand the two competing groups use Bayesian estimation methods implemented with their own software packages. The use of the same estimation method is very helpful in terms of making comparisons. The two crab team groups use completely different estimation methods, neither of which is entirely satisfactory.

The weighted least squares method (Zheng 2004) does not allow the production of standardized residuals. It is not a "fully statistical" model: diagnostics cannot be properly evaluated. The maximum likelihood approach (Turnock & Rugolo 2005) at least allows the production of standardized residuals even if this has not been routinely done. However, the use of penalty functions is not ideal and where possible they should be replaced by properly formed priors. Indeed, both modeling groups need to move towards fully Bayesian assessment methods as soon as possible. There simply is no other generally accepted method for incorporating prior/ancillary information and statistically accounting for uncertainty. It is not perfect, but it is currently the "state of the art" and will be so for some time to come.

The assumption by both modeling groups that the trawl survey q is known exactly (on the basis of "under-bag" experiments) ignores the uncertainty due to unknown aerial availability (i.e., the proportion of the population within the survey area). The estimates are conservative but they are definitely biased. The assumption is neither necessary nor desirable and the trawl survey q should be estimated. The information from the under-bag experiments, and whatever else is "known" about q, can be incorporated in a prior (or, if necessary, a penalty function).

The estimation of natural mortality (M) is always problematic whether it is done inside or outside of a stock assessment model. This is true for a single M assumed constant over the whole history of a fishery. Attempts to estimate different M for different time periods in a stock assessment (Zheng 2004) are ill-advised unless there are ancillary data which can reasonably be argued to index M in some way (e.g., a biomass index for a known major predator).

Many of the problems facing the crab team are generic in nature, some are crab specific (definition of B) and some are even more general (testing of OFDs). Wherever possible, efforts should be made to establish collaborative projects to share the workload.

CONCLUSIONS

The conclusions are organized according to the headings provided in the SOW (Appendix 3).

a. Strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.

The proposed OFD is:

- an improvement on the existing OFD
- comprehensive (as a framework)
- borrowed from groundfish (so is already reviewed to some extent).

The existing OFD does not provide any sensible constraints on fishing mortality and in that regard it appears fatally flawed. The proposed OFD will at least provide constraints.

Weaknesses of the proposed OFD:

- complicated
- it is still a work in progress
    - default values for parameters are not yet determined
    - sensible definition of B not determined/specified
    - criteria for determining optimal default parameters not determined/specified
- extensive simulations are needed to determine suitable default parameters
- it may potentially unnecessarily constrain existing harvest strategies.

The Review Panel were shown the results of preliminary simulations aimed at determining suitable default values for alpha, beta, and gamma. One can envisage an extensive suite of simulations which could determine suitable default values, but this can only happen after:

- sensible definitions of B are derived (being proportional to total fertilized egg production)
- the criteria for optimal default parameter values are defined.

An important issue, relating to the optimality of default parameter values, is how to define a "good OFD". The current simulations test an OFD MSY control rule by using it as a HS (i.e., assuming that catch is always set at the OFL) and testing its performance when the stock is initially overfished. However, this ignores the fact that a council is required to act in a precautionary manner when setting TACs and, as such, the simulations are testing something which will never occur.

b. Recommendations for improvements to proposed overfishing definitions or alternative definitions.

The parameterization of the proposed MSY control rules implies a reduction in F at B_{MSY}. It does allow flat control rules (alpha = –infinity) but it precludes the suggested default overfishing definitions of Restrepo et al. 1998 (where the reduction in F occurs below B_{MSY}). I suggest that an extra parameter is added to the framework to allow MSY control rules of the form proposed by Restrepo et al. (1998). In the absence of this parameter, the proposed framework may unnecessarily restrict harvest strategies.

c. Review of model configurations, formulations and methods used to account for uncertainty.

The model configurations and formulations are generally appropriate, but there may have been some implementation error in some of the models (e.g., mating dynamics not consistent with expert opinion). There needs to be more effort made to ensure that both modeling groups correctly implement the agreed population dynamics. When alternative dynamics are considered possible they should also be implemented to allow sensitivity analyses to be performed.

d. Review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in the simulation models.

The determination of an appropriate SRR is one of the central issues of the review. All of the alternative definitions of B used in the preliminary simulations were inappropriate. They were either demonstrably inadequate (e.g., female mature biomass, total mature biomass) or inadequately justified (i.e., no analysis or derivation). Total fertilized egg production does not appear to be proportional to female mature biomass. Therefore, definitions of B should not be based on scaled female mature biomass (e.g., through an assumed mating ratio) .

Other life history parameters appeared to be appropriately estimated (except, in one model where M was estimated to change during different time periods – not appropriate without additional data – see recommendations).

e. Research priorities to improve understanding of essential population and fishery dynamics necessary to formulate best management practices.

Recommendations, with regard to all aspects of the review, are given in the section below.

RECOMMENDATIONS

It appears that the proposed overfishing framework can be considered "acceptable" (complete) without default parameter values, F_{MSY} proxies, or a definition for B. If this is the case, then the first assessment for each stock, under the new framework, will require a full suite of simulation results justifying the OFD used. As more stocks are assessed "default" definitions and parameter values will materialize as scientists borrow them from the previously accepted assessments. Such an evolution is far from ideal and the process will need to be managed carefully. It would be better to get agreement on as much as possible in the proposed OFD before it is "accepted". Certainly a default definition for B is desirable.

In any case, irrespective of the timing relative to the acceptance of the proposed OFD framework, I have the following recommendations.

- Derive sensible definitions of B:
    - being defensibly proportional to total fertilized egg production
    - consider primiporous and multiporous matings separately
    - get the cycle consistent with the best available expert opinion (i.e., which males can participate in which matings)
    - B is not proportional to mature female biomass (e.g., clutch size is not proportional to biomass)
    - use an analytical approach to derive suitable functional relationships (see Appendix 2)
    - estimate parameters of the relationship in the stock assessment models using available data on egg production by color class (see Appendix 2)
    - mature male biomass appears to be defensible (use as a default?)

- Agree on the criteria and method for testing (OFD) MSY control rules:
    - these methods could be applied to tiers 1-4 (e.g., not only to determine "good" alpha and beta values, but also to choose between different proxies, e.g., $F_{50\%}$ or $F_{60\%}$)
    - it must be decided what makes one MSY control rule better than another when they are part of an OFD (i.e., test their *function*, they are not rebuilding plans or harvest strategies)
    - test MSY control rules in conjunction with a HS (e.g., an existing HS or a "default" OY control rule which takes 75% of the OFL – see Restrepo et al. 1998)

- compare with a flat control rule ($F = F_{MSY}$, i.e. are alpha and/or beta even needed?)
- examine performance over a range of starting biomasses (not just overfished; you want to know how they perform "going down" as well as "going up")
- incorporate observation error (i.e., true B and observed B can differ)
- incorporate stochastic recruitment
- examine trade-off statistics (e.g., what is forgone in yield to achieve higher biomass/lower probability of being declared overfished)
- use the full definition of MSST (i.e., not just 0.5 B_{MSY} – see Restrepo et al. 1998)
- include an extra parameter in the MSY control rules so as not to exclude the suggested default rules of Restrepo et al. (1998) (this parameter can have a default of 0 if desired)

The following two recommendations only apply if it is decided to use tier 1-2 simulations to derive default alpha and beta for tiers 1-4. It may not be the case that "good" alpha and beta values in tiers 1-2 will necessarily be any good when used in conjunction with F_{MSY} proxies. However, it may be a necessary assumption given time constraints.

- Agree on criteria for testing F_{MSY} proxies (stock specific, Tier 3):
  - using expert judgment choose a range of steepness/SRR relationships (after sensible definition of B)
  - use minimax or some other agreed principle to choose the best proxy

- Agree on criteria for testing gamma (group specific, Tier 4):
  - explicitly and precisely define gamma (in relation to selectivity and timing of the fishery)
  - use the same approach as for tier 3, but wider parameter space
  - obtain default gamma for each of several species/stock groups

- Consider what simulations, if any, could help for tier 5:
  - to define the period over which catches should be averaged (e.g., guiding principles on "not too much catch variability"; not a "declining trend in biomass" over the period)

- Stock assessment models
  - estimate the survey catchability q
  - start with parsimonious models
  - only introduce extra parameters if absolutely necessary
  - do not confound M with possible changes in catchability
  - estimating changes in M is only defensible if supported by auxiliary data on known predators/disease
  - calculate standardized residuals
  - iteratively re-weight indices so that residuals are consistent with variance assumptions

    - as soon as possible move to fully Bayesian assessments

- Trawl survey

    - if feasible, routinely retain a sample of female crabs with orange colored eggs to estimate the proportion of fertilized orange-colored egg-production (i.e. to estimate, at the time of the survey, what proportion of orange colored eggs are actually fertilized)
    - if feasible, routinely retain a sample of females (of the relevant species) to estimate “sperm load” (i.e., for those species which retain sperm).

REFERENCES

(see Appendix 1 for further references)

Mace, P.M.; Doonan, I.J. 1988: A generalised bioeconomic simulation model for fish population dynamics. New Zealand Fisheries Assessment Research Document 88/4. 51 p.

Starr, P.; Annala, J.H.; Hilborn, R. 1998: Contested stock assessment: two case studies. *Can. J. Fish. Aquat. Sci. 55*: 529–537.

APPENDIX 1: MATERIAL PROVIDED

The website from the interagency workshop was made available to the Review Panel. This contained documents and presentations, but also contained links to other related material. Below I list the material which I obtained from the website (or related links) and additional documents which were emailed to the Review Panel or provided as hardcopy, before or during the review meeting. I do not include several documents which were emailed to the Review Panel after the meeting (as I did not consult them).

Anon. 1998. Fishery Management Plan for Bering Sea/Aleutian Islands King and Tanner Crabs. Exec summary. July 18, 1998. 5 p.

Anon. 1999. Draft for Secretarial Review: Environmental Assessment for Amendment 7 to the Fishery Management Plan for the commercial king and tanner crab fisheries in the Bering Sea/Aleutian Islands. 53 p.

Anon. 2005. Magnuson-Stevens Act Provisions; National Standard Guidelines; Proposed Rule. Federal Register Vol 70, No. 119. 21 p.

Anon. 2006. Center of Independent Experts, Alaska Crab Overfishing Definitions, April 24-29, 2006. Alaska Fisheries Science Center, Seattle, WA. Apr 14th 2006 Draft Agenda. 2 p.

Anon. 2006. Current Overfishing Definitions in Crab FMP (FMP Section 6.0 revised from Amendment 7 1998). 2 p.

Anon. 2006. Workshop Report Crab Overfishing Definitions Inter-agency Workshop, February 28-March 1, 2006. Alaska Fisheries Science Center Seattle, WA. 21 p.

Anon. 2006. Alaska Crab Overfishing Definitions Workshop, February 28 – March 1, 2006. Alaska Fisheries Science Center, Seattle, WA. Feb 22, 2005. Draft Agenda . 2 p.

Anon. 2006. Draft report of the Scientific and Statistical Committee to the North Pacific Fishery Management Council, April 3-5, 2006 . 17 p.

Anon. 2006. Participant list for interagency workshop. 1 p.

Anon. 2006. Roadmap for Crab Workshop Documents. 1 p.

Anon. 2006. Tier system. 7 slides.

Maunder, M.N. 2003. Review of the stock assessment and harvest strategy for the eastern Bering Sea snow crab. CIE review report. 29 p.

Punt, A.E. 2003. The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. *Fisheries Research 65*: 391–409.

Restrepo, V.R. et al. 1998: Technical guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act. 56 p.

Rugolo, L.J.: Simplifications in population models and fishery dynamics modeling approaches. 16 slides

Rugolo, L.J.: Statement of work: NPFMC BSAI King & Tanner Crab Working Group. 30 slides.

Rugolo, L.J.; Siddeek, M.S.M., Turnock, B.J.; Zheng, J. 2004. North Pacific Fisheries Management Council Bering Sea / Aleutian Islands King and Tanner Crab Working Group Progress Report to the Crab Plan Team 22 September 2004. 34 p.

Siddeek, M.S.M: Parameters input to SPR models. 8 slides.
Siddeek, M.S.M.: Preliminary results. 7 slides.
Siddeek, M.S.M.: Model structures. 3 slides + 2 spreadsheets.
Siddeek, M.S.M.: Approaches to estimate proxy BRP values. 9 slides.
Siddeek, M.S.M.; Zheng, J. 2006. Reference point estimation analysis for the Bering Sea and Aleutian Islands (king and tanner) crab revised Fisheries Management Plan. 68 p.
Stram, D.L.: Overview of crab management and background on current overfishing definitions. Alaska Crab Overfishing Definitions Workshop, AFSC, Seattle, WA, February 28-March 1, 2006. 10 slides.
Thompson, G.: National Standard 1 Guidelines: use of SPR reference points, and incorporating uncertainty. 14 slides.Turnock, B.J. Snow crab stock assessment. 30 slides
Turnock, B.J.: Proposed tier system. 9 slides.
Turnock, B.J.: Input values to SPR models estimated from stock assessment models. 15 slides.
Turnock, B.J.; Rugolo, L.J. 2005. Stock Assessment of eastern Bering Sea snow crab. 96 p.
Turnock, B.J.; Rugolo, L.J. 2006. Analysis of proposed overfishing tier system for BSAI king and tanner crab stocks. 28 p.
Zheng, J.: Population Dynamics and Stock Assessment of Red King Crab in Bristol Bay, Alaska. 29 slides
Zheng, J. 2004. Bristol Bay red king crab stock assessment in 2004. 72 p.

APPENDIX 2: DEFINITION OF B IN THE SRR

Below I present three suggestions for the definition of B in the SRR, each being a proxy for total fertilized egg production (TFEP). They range from simple to complex. For the most complex method I also illustrate how some of the unknown parameters in the functional form might be estimated.

The three suggested definitions for B in the SRR are:

- total mature male biomass
- a function of total mature male biomass
- a function of total female egg production and a "fertilization factor".

The first suggestion was made by Dr Caputi in the review meeting. It is appealing in its simplicity and it makes sense for crab stocks when the male population is severely depleted. It is deficient for crab populations when the sex ratio is near its virgin level. However, we can derive a functional form to correct for its deficiency.

Suppose that,

$$TFEP = \min \{ P, q\, SP \}$$

where P = total egg production, SP = sperm production, and q is a constant which translates sperm production into number of eggs. In reality, q is not a constant; it depends on any and all factors which affect fertilization success (e.g., size and age structure, environmental effects on sperm potency, spawning migration patterns). In fish stocks we expect that $q\, SP >> P$ and thus accept any reasonable proxy for P as a proxy for TFEP for use in the SRR. However, in crab stocks where fishing mortality is directed only at males we must incorporate sperm production.

Consider a deterministic model for a crab population where fishing mortality (F) is primarily on males. Let,

B_F = male mature biomass at equilibrium under fishing mortality F
P_F = total egg production at equilibrium under fishing mortality F
$a_F = P_F / B_F$

and let $TFEP_F$ and SP_F denote TFEP and SP respectively at equilibrium under fishing mortality F.

We then have,

$$TFEP_F = \min \{ P_F, q\, SP_F \}$$

Now, suppose that sperm production is proportional to male biomass: $SP_F = s\, B_F$.

Hence, when q $SP_F \leq P_F$, we have

$$TFEP_F = q\, s\, B_F$$

and when q $SP_F \geq P_F$, we have

$$TFEP_F = P_F = a_F\, B_F.$$

Note, that when q $SP_F = P_F$, we have a_F = q s. Denote the F at this point as F_a and the associated B as $B(F_a)$.

Now, as F varies from 0 to infinity, B_F varies from its virgin level, B_0, to 0. From 0 to $B(F_a)$, $TFEP_F$ is a linear function of B_F (which passes through the origin). The form of $TFEP_F$ between $B(F_a)$ and B_0 depends on the nature of a_F. However, since males are preferentially exploited it follows that as F decreases that a_F also decreases. Hence, $TFEP_F$ is a linear function of B_F from 0 to $B(F_a)$, and then is convex from $B(F_a)$ to B_0.

This deduction allows us to use an approximate functional form which is independent of the details of any particular model. We will use an exponential function which is approximately linear over part of its range and then convex. Changing the notation somewhat, let,

$$TFEP(B) = b\,[\,1 - \exp(\,-aB)\,]$$

where B = male mature biomass, and a, b, are unknown parameters.

Let, $TFEP(B_0) = P_0$, then since,

$$TFEP(B_0) = b\,[\,1 - \exp(\,-aB_0)\,] = P_0$$

it follows that,

$$a = \frac{-\ln(1-\frac{P_0}{b})}{B_0}$$

and

$$\text{TFEP}(B) = b\left[1-\left(1-\frac{P_0}{b}\right)^{\frac{B}{B_0}}\right]$$

This can be better expressed as,

$$\text{TFEP}(B) = \frac{P_0}{1-\eta}\left[1-\eta^{\frac{B}{B_0}}\right]$$

where $0 < \eta < 1$.

This equation provides a simple generalization of total male mature biomass as the definition of “B”. The range of η values to consider depends on how effective one believes the males can be at fertilizing eggs when at depleted biomass levels (or how many surplus males there were at virgin levels). Values of $\eta > 0.5$ provide fairly linear functions, with TFEP at 20% B_0 not much more than 20% P_0. For approximately 40% P_0 at 20% B_0 use $\eta = 0.1$ and for approximately 60% P_0 at 20% B_0 use $\eta = 0.01$.

This equation should be used in conjunction with an assumed SRR as a function of TFEP to produce a SRR as a function of mature male biomass. For example, if mean recruitment is given by R(TFEP), then use R(TFEP(B)) – i.e., to get mean recruitment as a function of B rather than TFEP.

For example, if a Beverton Holt SRR is assumed with steepness Δ (Mace & Doonan 1988), then

$$R = R_0 \left[\frac{4\Delta\left(1-\eta^{\frac{B}{B_0}}\right)}{(1-\Delta)(1-\eta)+(5\Delta-1)\left(1-\eta^{\frac{B}{B_0}}\right)} \right]$$

which behaves much like a Beverton Holt SRR unless both Δ and η are small (i.e., is similar to a Beverton Holt SRR with Δ set equal to the proportion of R_0 obtained at 20%B_0 from the full relationship). Information on η may be available from trawl survey data in which case it may be possible to estimate η within the stock assessment model, or even externally.

The third suggested definition for B is arrived at using a somewhat different approach. The idea is to start with an unspecified functional form and then, through a series of assumptions, and bearing in mind what data are available, arrive at a particular form.

Consider a particular mating (i.e., the primiporous or multiporous mating, or perhaps a combined mating for modeling convenience) in a particular year and suppose that sperm are not stored by the females (“sperm storage” across matings could perhaps be incorporated if data on stored sperm levels from the trawl survey were available).

Let $C = \{ c_i \mid i \, \varepsilon \, n \}$ be the set of all crabs involved in the mating (i.e., there are n crabs labeled 0,…,n –1). Each crab has various biological characteristics. E.g,: $s(c_i)$ = sex of the ith crab, $cw(c_i)$ = carapace width of the ith crab, $a(c_i)$ = age of ith crab, $f(c_i)$ = clutch fullness category of ith crab. Subsets of crabs can be specified. E.g., s(C, male) = $\{ c_i \mid s(c_i) = \text{male} \}$. Let TFEP = $\Gamma(C)$. That is, the total fertilized egg production (of this mating) potentially depends on every characteristic of every crab involved in the mating. True, but unhelpful. We will split TFEP into tow components: total egg production (P), and a fertilization factor (G):

$$\Gamma(C) = P[\, s(C, fem)\,]\, G(C)$$

Now, we have assumed that egg production depends only on the females; fertilization is still dependent on all crabs (e.g., male and female length frequencies, as well as sex ratio).

From data presented during the review meeting and subsequent discussion it is clear that quite a lot is known about clutch fullness as a function of various categorical variables (e.g., primiporous females typically have 0.75 clutch fullness, 2nd clutches are typically full, shell condition 4 and 5 females are typically barren, older females have lower clutch fullness). For a given clutch fullness it appeared that egg production of an individual female was a linear function of carapace width. In any case, the data exist and can be analyzed to provide an appropriate functional form and parameterization for P. One approach would be to use a GLM to explain individual egg production as a function of variables which could reasonably be incorporated in the stock assessment model.

For example, for a combined mating, the data may be consistent with a single categorization within which a linear function of carapace width may be adequate for average individual egg production. The functional form could be as follows:

$$P[s(C, fem)] = \sum_{i \varepsilon cat} \sum_{j \varepsilon cat_i} a_i + b_i cw(c_j)$$

where cat = { primiporous, 2nd mating, shell condition > 3, other }, a_i and b_i are the linear coefficients for each category member, and cw denotes carapace width.

The above form is just an example which may or not be suitable. However, given the available data I am confident that a suitable form will be derived. It will be defensible and I doubt that female mature biomass will be seen to be an adequate proxy.

The fertilization factor is a more difficult challenge. However, there are also data available which may enable the estimation of relevant parameters if an appropriate form can be hypothesized. The simplest form for G(C) is a constant. However, this would make TFEP independent of males – clearly not appropriate. The fertilization factor, at a minimum, must use the relative number of males and females. Other elements, such as relative size distributions and the propensity for males to fight for “desirable females” could also be brought in (but not easily).

A candidate, already used as a component of some of the Work Group definitions of B is:

G(C) = min { n[s(C, male)] / n[s(C, fem)] r, 1 }

where r is an unknown “mating ratio” and n[] denotes cardinality of a set (i.e., the number of members).

There are data available from the trawl surveys which could be used to estimate r, preferably within the stock assessment model (so that trawl survey catchability and selectivity can be estimated simultaneously). I refer to the individual egg production data which includes a color classification. Clutches are orange to begin with and at the time of the survey clutches are either orange or another color. If they are non-orange then they are fertilized, but some proportion of orange clutches are also fertilized (and haven't changed color yet). To use these data in a stock assessment model, we need to be able to formulate predictions for orange and non-orange trawl-survey egg production. Below I give a sketch of how to do this.

We already have an expression for total egg production, P[s(C, fem)] which could be modified in the model to account for trawl survey selectivity and catchability to become an expression for "trawl-survey egg production". We shall denote this simply by P. In the following, assume that trawl-survey selectivity and catchability have been appropriately dealt with in all components of (predicted) egg production.

Let,

P_o = orange egg production
P_{of} = orange fertilized egg production
P_{oun} = orange unfertilized egg production
P_{non} = non-orange egg production
p_f = proportion of fertilized eggs
p_{fo} = proportion of fertilized eggs that are orange

We have,

$P_o = P_{of} + P_{oun} = P\, p_f\, p_{fo} + P\,(1 - p_f)$

and

$P_{non} = P - P_o$.

Which gives,

$P_0 = P\,[\, p_f\, p_{fo} + 1 - p_f \,]$

$P_{non} = P\, p_f\,[\, 1 - p_{fo} \,]$.

We have two unknown parameters: p_f and p_{fo}.

However, p_f = n[s(C, male)] / n[s(C, fem)] r. So there is only one extra parameter: p_{fo}. Data on this could be available for each (future) survey if a random sample of females with orange clutches was retained and observed in the lab to see what proportion of orange clutches were fertilized (noting that p_{fo} = "proportion orange and fertilized"/p_f).

Alternatively, some assumptions about the distribution of mating times would be needed together with some knowledge of how long it takes fertilized eggs to change color.

For example, assume a normal distribution for the mating time X of a female: $X \sim \mathrm{N}(t_m, \sigma_m^2)$. Suppose that a females' clutch will change color after a time interval δ, and let Y = X + δ. Suppose that the survey occurs at time t_s and let $q(t_s)$ = the proportion of (fertilized) eggs that are non-orange.

Then,

$$q(t_s) = \mathrm{Prob}\left(Y < t_s\right) = \mathrm{Prob}\left(Z < \frac{t_s - (t_m + \delta)}{\sigma_m} \right)$$

where Z is the standard normal random variable. Some educated guesses will help define a range for $q(t_s)$ and hence to $p_{fo} = 1 - q(t_s)$. Another approach is to look at the relative distribution of clutches within color classes to try to directly estimate the proportion of orange clutches which are fertilized (e.g., a disjunction between the proportion of orange clutches and the proportion of the next color class may indicate that very few orange clutches are fertilized – for a particular survey).

Trying to include competition between males is an interesting exercise. It can be approached by setting up a system of differential equations with coupling, decoupling, and fighting rates. It gets sufficiently complicated that it may not be a worthwhile exercise in itself. Perhaps it is better to do the full job and look to set up a system of differential equations for crab-specific population dynamics – a medium to long term project.

APPENDIX 3: STATEMENT OF WORK

The statement of work given below was received in early May after I returned from the Seattle meeting. It differs from the original statement of work in two respects. First, it clarified some issues which the Review Panel raised while we were in Seattle. Second, it contains a new date for submission of reports. I was not able to accommodate the shift of the deadline from 1 June 2006 to 12 May 2006. However, I did produce an interim report with the highlights of my findings and recommendations which I supplied to Dr Bell before his attendance at the May 17 meeting.

Consulting Agreement between the University of Miami and Reviewer

April 27, 2006

Background

The Alaska Fisheries Science Center (AFSC) requests review of proposed overfishing definitions and simulation models used to evaluate biological reference points for Bering Sea and Aleutian Islands King and Tanner crab stocks. The North Pacific Fishery Management Council (NPFMC) has determined that the existing overfishing definitions for Bering Sea and Aleutian Islands King and Tanner crab stocks need revision. The AFSC is seeking review of the population dynamics models developed for revising the overfishing definitions.

There are currently 22 Bering Sea and Aleutian Islands crab stocks under the Federal Bering Sea Aleutian Island Crab Fishery Management Plan (FMP) of which 7 are considered major stocks. Four of the seven major crab stocks have been declared overfished and rebuilding plans developed within the last 7 years. Of the remaining three stocks, only one has been relatively stable at a low level, another has maintained stable catch for several years, however, even for this stock it appears recruitment may be declining. While the remaining stock has increased, survey abundance estimates have low precision and the fishery is closed due to bycatch concerns. There is no consensus on the principal cause of declines in Bering Sea crab stocks.

Review Requirements

A panel of three consultants is requested for this review. In aggregate, the panel will need to be thoroughly familiar with various subject areas involved in the review: crab biology; analytical stock assessment, including population dynamics theory, length-based stock assessment models, rebuilding analyses, estimation of biological reference points and harvest strategy modeling for invertebrates; and AD Model Builder. The CIE consultants will travel to Seattle, Washington to meet with the Interagency Work Group charged with developing the new overfishing definitions. We request that one member of the Panel should be present at the May meeting of the NPFMC Crab Plan Team in

Anchorage, Alaska. We also request that one member of the Panel be present at the June meeting of the NPFMC Scientific and Statistical Committee meeting in Kodiak, Alaska. It would be preferable that the same individual attends both of these meetings, but this is not a requirement.

The report generated by each consultant should include:

a. A statement of the strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.
b. Recommendations for improvements to proposed overfishing definitions or alternative definitions,
c. A review of the model configurations, formulations and methods used to account for uncertainty.
d. A review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in simulation models.
e. Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

AFSC will provide copies of the NPFMC Work Group statement of work, proposed overfishing definitions, preliminary results of simulations, discussion of input parameters, a copy of the code for the snow crab stock assessment, and the AD Model Builder and Fortran code used for reference point estimation. The panel will meet with scientists from the Alaska Fisheries Science Center and the Alaska Department of Fish and Game from April 24 to April 28, 2006, in Seattle, Washington (see attached agenda).

It is estimated that the duties of each reviewer will occupy a maximum of 14 days each: several days for preparation, five days for the workshop, several days for writing their reports, and two days for travel. In addition, a maximum of nine reviewer days will be allowed for attending the two council meetings, including preparation time, travel, and one day to attend each meeting. The total level of effort is 51 days of reviewer time.

Products

- One member of the panel will attend the May meeting of the Crab Plan Team on May 17, 2006 in Anchorage, Alaska, to discuss the panel's findings regarding the strengths and weaknesses of proposed definitions and modeling approaches.
- One member of the Panel will attend the June meeting of the NPFMC Scientific and Statistical Committee meeting on June 5, 2006 in Kodiak, Alaska, to discuss the panel's findings regarding the strengths and weaknesses of proposed definitions and modeling approaches.
- No later than May 12, 2006, each panelist shall submit a written report of findings, analysis, and conclusions. See Annex 1 for details on the report outline. The reports should be sent via e-mail to Dr. David Die at ddie@rsmas.miami.edu, and to Mr. Manoj Shivlani at mshivlani@rsmas.miami.edu.

ANNEX I: REPORT GENERATION AND PROCEDURAL ITEMS

1. The report should be prefaced with an executive summary of findings and/or recommendations.

2. The main body of the report should consist of a background, description of review activities, summary of findings, and conclusions/recommendations.

3. The report should also include as separate appendices the bibliography of materials provided by the Center for Independent Experts and the center and a copy of the statement of work.

4. Individuals shall be provided with an electronic version of a bibliography of background materials sent to all reviewers. Other material provided directly by the center must be added to the bibliography that can be returned as an appendix to the final report.

Please refer to the following website for additional information on report generation: http://www.rsmas.miami.edu/groups/cimas/Report_Standard_Format.html